

(12) INTERNATIONAL APPLICATION PUBLISHED UNDER THE PATENT COOPERATION TREATY (PCT)

(19) World Intellectual Property Organization  
International Bureau



(43) International Publication Date  
4 December 2003 (04.12.2003)

PCT

(10) International Publication Number  
**WO 03/099298 A1**

- (51) International Patent Classification<sup>7</sup>: **A61K 31/7105**, 31/711, 31/713 (74) Agents: **WEICKMANN & WEICKMANN** et al.; Postfach 860 820, 81635 München (DE).
- (21) International Application Number: **PCT/EP03/05513** (81) Designated States (*national*): AE, AG, AL, AM, AT, AU, AZ, BA, BB, BG, BR, BY, BZ, CA, CH, CN, CO, CR, CU, CZ, DE, DK, DM, DZ, EC, EE, ES, FI, GB, GD, GE, GH, GM, HR, HU, ID, IL, IN, IS, JP, KE, KG, KP, KR, KZ, LC, LK, LR, LS, LT, LU, LV, MA, MD, MG, MK, MN, MW, MX, MZ, NI, NO, NZ, OM, PH, PL, PT, RO, RU, SC, SD, SE, SG, SK, SL, TJ, TM, TN, TR, TT, TZ, UA, UG, US, UZ, VC, VN, YU, ZA, ZM, ZW.
- (22) International Filing Date: 26 May 2003 (26.05.2003)
- (25) Filing Language: English
- (26) Publication Language: English
- (30) Priority Data: 60/382,983 24 May 2002 (24.05.2002) US (84) Designated States (*regional*): ARIPO patent (GH, GM, KE, LS, MW, MZ, SD, SL, SZ, TZ, UG, ZM, ZW), Eurasian patent (AM, AZ, BY, KG, KZ, MD, RU, TJ, TM), European patent (AT, BE, BG, CH, CY, CZ, DE, DK, EE, ES, FI, FR, GB, GR, HU, IE, IT, LU, MC, NL, PT, RO, SE, SI, SK, TR), OAPI patent (BF, BJ, CF, CG, CI, CM, GA, GN, GQ, GW, ML, MR, NE, SN, TD, TG).
- (71) Applicant (*for all designated States except US*): **MAX-PLANCK GESELLSCHAFT ZUR FÖRDERUNG DER WISSENSCHAFTEN E.V.** [DE/DE]; Hofgartenstrasse 8, 80539 München (DE).
- (72) Inventors; and
- (75) Inventors/Applicants (*for US only*): **TUSCHL, Thomas** [DE/DE]; Keplerstrasse 9, 37085 Göttingen (DE). **EL-BASHIR, Sayda** [DE/DE]; Ludwig-Beck-Str. 17, 37075 Göttingen (DE). **LENDECKEL, Winfried** [DE/DE]; Blinde Gasse 52, 37318 Hohengandern (DE). **WEST-PHAL, Christoph** [US/US]; 12 Willard St., Newton, MA 02458 (US).
- Published:**
- with international search report
  - before the expiration of the time limit for amending the claims and to be republished in the event of receipt of amendments
- For two-letter codes and other abbreviations, refer to the "Guidance Notes on Codes and Abbreviations" appearing at the beginning of each regular issue of the PCT Gazette.*



**WO 03/099298 A1**

(54) Title: RNA INTERFERENCE MEDIATING SMALL RNA MOLECULES

(57) Abstract: Double-stranded RNA (dsRNA) induces sequence-specific post-transcriptional gene silencing in many organisms by a process known as RNA interference (RNAi). Using a *Drosophila* in vitro system, we demonstrate that 19-23 nt short RNA fragments are the sequence-specific mediators of RNAi. The short interfering RNAs (siRNAs) are generated by an RNase III-like processing reaction from long dsRNA. Chemically synthesized siRNA duplexes with overhanging 3' ends mediate efficient target RNA cleavage in the lysate, and the cleavage site is located near the center of the region spanned by the guiding siRNA. Furthermore, we provide evidence that the direction of dsRNA processing determines whether sense or antisense target RNA can be cleaved by the produced siRNP complex.

## RNA INTERFERENCE MEDIATING SMALL RNA MOLECULES

### RELATED APPLICATIONS

This application claims priority to U.S. provisional application number 60/382,983  
5 filed on May 24, 2002, the contents of which are incorporated herein by reference.

### BACKGROUND

The term "RNA interference" (RNAi) was coined after the discovery that injection of dsRNA into the nematode *C. elegans* leads to specific silencing of genes highly homologous  
10 in sequence to the delivered dsRNA (Fire *et al.*, 1998). RNAi was subsequently also observed in insects, frogs (Oelgeschlaeger *et al.*, 2000), and other animals including mice (Svoboda *et al.*, 2000; Wianny and Zernicka-Goetz., 2000) and is likely to also exist in humans. RNAi is closely linked to the post-transcriptional gene-silencing (PTGS) mechanism of co-suppression in plants and quelling in fungi (Catalanotto *et al.*, 2000;  
15 Cogoni and Macino, 1999; Dalmay *et al.*, 2000; Ketting and Plasterk, 2000; Mourrain *et al.*, 2000; Smardon *et al.*, 2000) and some components of the RNAi machinery are also necessary for post-transcriptional silencing by co-suppression (Catalanotto *et al.*, 2000; Demburg *et al.*, 2000; Ketting and Plasterk, 2000). The topic has also been reviewed recently (Bass, 2000; Boshier and Labouesse, 2000; Fire, 1999; Plasterk and Ketting, 2000; Sharp, 1999; Sijen and  
20 Kooter, 2000; Plant Molecular Biology, vol. 43, issue 2/3, 2000).

In plants, in addition to PTGS, introduced transgenes can also lead to transcriptional gene silencing via RNA-directed DNA methylation of cytosines (see references in Wassenegger, 2000). Genomic targets as short as 30 bp are methylated in plants in an RNA-directed manner (Pelissier, 2000). DNA methylation is also present in mammals.

25 The natural function of RNAi and co-suppression appears to be protection of the genome against invasion by mobile genetic elements such as retrotransposons and viruses which produce aberrant RNA or dsRNA in the host cell when they become active (Jensen *et al.*, 1999; Ketting *et al.*, 1999; Ratcliff *et al.*, 1999; Tabara *et al.*, 1999). Specific mRNA degradation prevents transposon and virus replication although some viruses are able to  
30 overcome or prevent this process by expressing proteins that suppress PTGS (Lucy *et al.*, 2000; Voimmet *et al.*, 2000).

DsRNA triggers the specific degradation of homologous RNAs only within the region of identity with the dsRNA (Zamore *et al.*, 2000). The dsRNA is processed to 21-23 nt RNA fragments and the target RNA cleavage sites are regularly spaced 21-23 nt apart. It has therefore been suggested that the 21-23 nt fragments are the guide RNAs for target recognition (Zamore *et al.*, 2000). These short RNAs were also detected in extracts prepared from *D. melanogaster* Schneider 2 cells which were transfected with dsRNA prior to cell lysis (Hammond *et al.*, 2000), however, the fractions that displayed sequence specific nuclease activity also contained a large fraction of residual dsRNA. The role of the 21-23 nt fragments in guiding mRNA cleavage is further supported by the observation that 21-23 nt fragments isolated from processed dsRNA are able, to some extent, to mediate specific mRNA degradation (Zamore *et al.*, 2000). RNA molecules of similar size also accumulate in plant tissue that exhibits PTGS (Hamilton and Baulcombe, 1999).

Here, we use the established *Drosophila in vitro* system (Tuschl *et al.*, 1999; Zamore *et al.*, 2000) to further explore the mechanism of RNAi. We demonstrate that short 21 and 22 nt RNAs, when base-paired with 3' overhanging ends, act as the guide RNAs for sequence specific mRNA degradation. Short 30 bp dsRNAs are unable to mediate RNAi in this system because they are no longer processed to 21 and 22 nt RNAs. Furthermore, we defined the target RNA cleavage sites relative to the 21 and 22 nt short interfering RNAs (siRNAs) and provide evidence that the direction of dsRNA processing determines whether a sense or an antisense target RNA can be cleaved by the produced siRNP endonuclease complex. Further, the siRNAs may also be important tools for transcriptional modulating, *e.g.*, silencing of mammalian genes by guiding DNA methylation.

Further experiments in human *in vivo* cell culture systems (HeLa cells) showed that double stranded RNA molecules having a length of preferably from 19-25 nucleotides have RNAi activity. Thus, in contrast to the results from *Drosophila* also 24 and 25 nt long double stranded RNA molecules are efficient for RNAi.

## SUMMARY

The present invention provides methods of mediating target-specific RNA interference. The compositions and methods described herein have improved efficacy and safety compared to prior art compositions and methods.

The invention further relates to the use of isolated double stranded RNA molecules, wherein each RNA strand has a length from 19-25 nucleotides, for mediating target specific nucleic acid modifications, particularly RNA interference (RNAi), in mammalian cells, particularly in human cells. Preferably, each strand of the RNA molecule has a length from 20-22 nucleotides (or 20-25 nucleotides in mammalian cells), wherein the length of each strand may be the same or different. Preferably, the length of the 3'-overhang reaches from 1-3 nucleotides, wherein the length of the overhang may be the same or different for each strand. The RNA-strands preferably have 3'-hydroxyl groups. The 5'-terminus preferably includes a phosphate, diphosphate, triphosphate or hydroxyl group. The most effective dsRNAs are composed of two 21 nt strands which are paired such that 1-3 nt (particularly 2 nt) 3' overhangs are present on both ends of the dsRNA.

One aspect of the invention relates to a method of mediating target-specific nucleic acid modifications, particularly RNA interference and/or DNA methylation, in a cell or an organism. The method can include the following steps:

- (a) contacting the cell or organism with the double stranded RNA molecule of the invention under conditions wherein target-specific nucleic acid modifications may occur and
- (b) mediating a target-specific nucleic acid modification effected by the double stranded RNA towards a target nucleic acid having a sequence portion substantially corresponding to the double stranded RNA.

The target gene to which the RNA molecule of the invention is directed may be associated with a pathological condition. For example, the gene may be a pathogen-associated gene (*e.g.*, a viral gene), a tumor-associated gene, or an autoimmune disease-associated gene. The target gene may also be a heterologous gene expressed in a recombinant cell or a genetically altered organism. By modulating, and particularly inhibiting, the function of such a gene, valuable information and therapeutic benefits in the agricultural and medical fields can be obtained.

In one aspect, the invention features a method of treating a subject at risk for or afflicted with unwanted cell proliferation, *e.g.*, malignant or nonmalignant cell proliferation. The method includes:

providing an siRNA, *e.g.*, an siRNA having a structure described herein, where the siRNA is homologous to and can silence, *e.g.*, by cleavage, a gene that promotes unwanted cell proliferation; and

5 administering the siRNA to a subject, preferably a human subject, thereby treating the subject.

In a preferred embodiment, the gene is a growth factor or growth factor receptor gene, a kinase, *e.g.*, a protein tyrosine, serine or threonine kinase gene, an adaptor protein gene, a gene encoding a G protein superfamily molecule, or a gene encoding a transcription factor.

10 In a preferred embodiment, the siRNA silences the PDGF beta gene, and thus can be used to treat a subject having or at risk for a disorder characterized by unwanted PDGF beta expression, *e.g.*, testicular and lung cancers.

In another preferred embodiment the siRNA silences the Erb-B gene, and thus can be used to treat a subject having or at risk for a disorder characterized by unwanted Erb-B expression, *e.g.*, breast cancer.

15 In a preferred embodiment, the siRNA silences the, Src gene, and thus can be used to treat a subject having or at risk for a disorder characterized by unwanted Src expression, *e.g.*, colon cancers.

20 In a preferred embodiment, the siRNA silences the CRK gene, and thus can be used to treat a subject having or at risk for a disorder characterized by unwanted CRK expression, *e.g.*, colon and lung cancers.

In a preferred embodiment, the siRNA silences the GRB2 gene, and thus can be used to treat a subject having or at risk for a disorder characterized by unwanted GRB2 expression, *e.g.*, squamous cell carcinoma.

25 In another preferred embodiment the siRNA silences the RAS gene, and thus can be used to treat a subject having or at risk for a disorder characterized by unwanted RAS expression; *e.g.*, pancreatic, colon and lung cancers, and chronic leukemia.

In another preferred embodiment the siRNA silences the MEKK gene, and thus can be used to treat a subject having or at risk for a disorder characterized by unwanted MEKK expression, *e.g.*, squamous cell carcinoma, melanoma or leukemia.

In another preferred embodiment the siRNA silences the JNK gene, and thus can be used to treat a subject having or at risk for a disorder characterized by unwanted JNK expression, *e.g.*, pancreatic or breast cancers.

In a preferred embodiment, the siRNA silences the RAF gene, and thus can be used to treat a subject having or at risk for a disorder characterized by unwanted RAF expression, *e.g.*, lung cancer or leukemia.

In a preferred embodiment, the siRNA silences the Erkl/2 gene, and thus can be used to treat a subject having or at risk for a disorder characterized by unwanted Erkl/2 expression, *e.g.*, lung cancer.

In another preferred embodiment the siRNA silences the PCNA(p21)gene, and thus can be used to treat a subject having or at risk for a disorder characterized by unwanted PCNA expression, *e.g.*, lung cancer.

In a preferred embodiment, the siRNA silences the MYB gene, and thus can be used to treat a subject having or at risk for a disorder characterized by unwanted MYB expression, *e.g.*, colon cancer or chronic myelogenous leukemia.

In a preferred embodiment, the siRNA silences the c-MYC gene, and thus can be used to treat a subject having or at risk for a disorder characterized by unwanted c-MYC expression, *e.g.*, Burkitt's lymphoma or neuroblastoma.

In another preferred embodiment the siRNA silences the JUN gene, and thus can be used to treat a subject having or at risk for a disorder characterized by unwanted JUN expression, *e.g.*, ovarian, prostate or breast cancers.

In another preferred embodiment the siRNA silences the FOS gene, and thus can be used to treat a subject having or at risk for a disorder characterized, by unwanted FOS expression, *e.g.*, skin or prostate cancers.

In a preferred embodiment, the siRNA silences the BCL-2 gene, and thus can be used to treat a subject having or at risk for a disorder characterized by unwanted BCL-2 expression, *e.g.*, lung or prostate cancers or Non-Hodgkin lymphoma.

In a preferred embodiment, the siRNA silences the Cyclin D gene, and thus can be used to treat a subject having or at risk for a disorder characterized by unwanted Cyclin D expression, *e.g.*, esophageal and colon cancers.

In a preferred embodiment, the siRNA silences the VEGF gene, and thus can be used to treat a subject having or at risk for a disorder characterized by unwanted VEGF expression, *e.g.*, esophageal and colon cancers.

5 In a preferred embodiment, the siRNA silences the EGFR gene, and thus can be used to treat a subject having or at risk for a disorder characterized by unwanted EGFR expression, *e.g.*, breast cancer.

In another preferred embodiment the siRNA silences the Cyclin A gene, and thus can be used to treat a subject having or at risk for a disorder characterized by unwanted Cyclin A expression, *e.g.*, lung and cervical cancers.

10 In another preferred embodiment the siRNA silences the Cyclin E gene, and thus can be used to treat a subject having or at risk for a disorder characterized by unwanted Cyclin E expression, *e.g.*, lung and breast cancers.

In another preferred embodiment the siRNA silences the WNT-1 gene, and thus can be used to treat a subject having or at risk for a disorder characterized by unwanted WNT-1 expression, *e.g.*, basal cell carcinoma.

15 In another preferred embodiment the siRNA silences the beta-catenin gene, and thus can be used to treat a subject having or at risk for a disorder characterized by unwanted beta-catenin expression, *e.g.*, adenocarcinoma or hepatocellular carcinoma.

20 In another preferred embodiment the siRNA silences the c-MET gene, and thus can be used to treat a subject having or at risk for a disorder characterized by unwanted c-MET expression, *e.g.*, hepatocellular carcinoma.

In another preferred embodiment the siRNA silences the PKC gene, and thus can be used to treat a subject having or at risk for a disorder characterized by unwanted PKC expression, *e.g.*, breast cancer.

25 In a preferred embodiment, the siRNA silences the NFkB gene, and thus can be used to treat a subject having or at risk for a disorder characterized by unwanted NFkB expression, *e.g.*, breast cancer.

In a preferred embodiment, the siRNA silences the STAT3 gene, and thus can be used to treat a subject having or at risk for a disorder characterized by unwanted STAT3 expression, *e.g.*, prostate cancer.

30

In another preferred embodiment the siRNA silences the survivin gene, and thus can be used to treat a subject having or at risk for a disorder characterized by unwanted survivin expression, *e.g.*, cervical or pancreatic cancers.

5 In another preferred embodiment the siRNA silences the Her2/Neu gene, and thus can be used to treat a subject having or at risk for a disorder characterized by unwanted Her2/Neu expression, *e.g.*, breast cancer.

In another preferred embodiment the siRNA silences the topoisomerase I gene, and thus can be used to treat a subject having or at risk for a disorder characterized by unwanted topoisomerase I expression, *e.g.*, ovarian and colon cancers.

10 In a preferred embodiment, the siRNA silences the topoisomerase II alpha gene, and thus can be used to treat a subject having or at risk for a disorder characterized by unwanted topoisomerase II expression, *e.g.*, breast and colon cancers.

In a preferred embodiment, the siRNA silences mutations in the p73 gene, and thus can be used to treat a subject having or at risk for a disorder characterized by unwanted p73 expression, *e.g.*, colorectal adenocarcinoma.

15 In a preferred embodiment, the siRNA silences mutations in the p21 (WAF1/CIP1) gene, and thus can be used to treat a subject having or at risk for a disorder characterized by unwanted p21 (WAF1/CIP1) expression, *e.g.*, liver cancer.

20 In a preferred embodiment, the siRNA silences mutations in the p27 (KIP1) gene, and thus can be used to treat a subject having or at risk for a disorder characterized by unwanted p27(KIP1) expression, *e.g.*, liver cancer.

In a preferred embodiment, the siRNA silences mutations in the PPM1D gene, and thus can be used to treat a subject having or at risk for a disorder characterized by unwanted PPM1D expression, *e.g.*, breast cancer.

25 In a preferred embodiment, the siRNA silences mutations in the RAS gene, and thus can be used to treat a subject having or at risk for a disorder characterized by unwanted RAS expression, *e.g.*, breast cancer.

30 In another preferred embodiment the siRNA silences mutations in the caveolin I gene, and thus can be used to treat a subject having or at risk for a disorder characterized by unwanted caveolin I expression, *e.g.*, esophageal squamous cell carcinoma.



In another preferred embodiment the siRNA silences mutations in the MIB1 gene and thus can be used to treat a subject having or at risk for a disorder characterized by unwanted MIB1 expression, *e.g.*, male breast carcinoma (MBC).

5 In another preferred embodiment the siRNA silences mutations in the MTA1 gene, and thus can be used to treat a subject having or at risk for a disorder characterized by unwanted MTA1 expression, *e.g.*, ovarian carcinoma.

In another preferred embodiment the siRNA silences mutations in the M68 gene, and thus can be used to treat a subject having or at risk for a disorder characterized by unwanted M68 expression, *e.g.*, human adenocarcinomas of the esophagus, stomach, colon, and  
10 rectum.

In preferred embodiments the siRNA silences mutations in tumor suppressor genes, and thus can be used as a method to promote apoptotic activity in combination with chemotherapeutics.

In a preferred embodiment, the siRNA silences mutations in the p53 tumor suppressor  
15 gene, and thus can be used to treat a subject having or at risk for a disorder characterized by unwanted p53 expression, *e.g.*, gall bladder, pancreatic and lung cancers.

In a preferred embodiment, the siRNA silences mutations in the p53 family member delta N-p63, and thus can be used to treat a subject having or at risk for a disorder characterized by unwanted delta N-p63 expression, *e.g.*, squamous cell carcinoma.

20 In a preferred embodiment, the siRNA silences mutations in the pRb tumor suppressor gene, and thus can be used to treat a subject having or at risk for a disorder characterized by unwanted pRb expression, *e.g.*, oral squamous cell carcinoma.

In a preferred embodiment, the siRNA silences mutations in the APC1 tumor suppressor gene, and thus can be used to treat a subject having or at risk for a disorder  
25 characterized by unwanted APC1 expression, *e.g.*, colon cancer.

In a preferred embodiment, the siRNA silences mutations in the BRCA1 tumor suppressor gene, and thus can be used to treat a subject having or at risk for a disorder characterized by unwanted BRCA1 expression, *e.g.*, breast cancer.

In a preferred embodiment, the siRNA silences mutations in the PTEN tumor  
30 suppressor gene, and thus can be used to treat a subject having or at risk for a disorder

characterized by unwanted PTEN expression; *e.g.*, hamartomas, gliomas, and prostate and endometrial cancers.

In a preferred embodiment, the siRNA silences MLL fusion genes, MLL2 *e.g.*, AF9, and thus can be used to treat a subject having or at risk for a disorder characterized by unwanted MLL fusion gene expression, *e.g.*, acute leukemias.

In another preferred embodiment the siRNA silences the BCR/ABL fusion gene, and thus can be used to treat a subject having or at risk for a disorder characterized by unwanted BCR/ABL fusion gene expression, *e.g.*, acute and chronic leukemias.

In another preferred embodiment the siRNA silences the TEL/AML1 fusion gene, and thus can be used to treat a subject having or at risk for a disorder characterized by unwanted TEL/AML1 fusion gene expression, *e.g.*, childhood acute leukemia.

In another preferred embodiment the siRNA silences the EWS/FLI1 fusion gene, and thus can be used to treat a subject having or at risk for a disorder characterized by unwanted EWS/FLI1 fusion gene expression, *e.g.*, Ewing Sarcoma.

In another preferred embodiment the siRNA silences the TLS/FUS1 fusion gene, and thus can be used to treat a subject having or at risk for a disorder characterized by unwanted TLS/FUS1 fusion gene expression, *e.g.*, Myxoid liposarcoma.

In another preferred embodiment the siRNA silences the PAX3/FKHR fusion gene, and thus can be used to treat a subject having or at risk for a disorder characterized by unwanted PAX3/FKHR fusion gene expression, *e.g.*, Myxoid liposarcoma.

In another preferred embodiment the siRNA silences the AML1/ETO fusion gene, and thus can be used to treat a subject having or at risk for a disorder characterized by unwanted AML1/ETO fusion gene expression, *e.g.*, acute leukemia.

In another aspect, the invention features, a method of treating a subject, *e.g.*, a human, at risk for or afflicted with a disease or disorder that may benefit by angiogenesis inhibition *e.g.*, cancer. The method includes:

providing an siRNA, *e.g.*, an siRNA having a structure described herein, which siRNA is homologous to and can silence, *e.g.*, by cleavage, a gene, which mediates angiogenesis; and

administering the siRNA to a subject, thereby treating the subject.

In a preferred embodiment, the siRNA silences the alpha v-integrin gene, and thus can be used to treat a subject having or at risk for a disorder characterized by unwanted alpha v-integrin, *e.g.*, brain tumors or tumors of epithelial origin.

5 In a preferred embodiment, the siRNA silences the Flt-1 receptor gene, and thus can be used to treat a subject having or at risk for a disorder characterized by unwanted Flt-1 receptors, *e.g.*, cancer and rheumatoid arthritis.

In a preferred embodiment, the siRNA silences the tubulin gene, and thus can be used to treat a subject having or at risk for a disorder characterized by unwanted tubulin, *e.g.*, cancer and retinal neovascularization.

10 In a preferred embodiment, the siRNA silences the tubulin gene, and thus can be used to treat a subject having or at risk for a disorder characterized by unwanted tubulin, *e.g.*, cancer and retinal neovascularization.

In another aspect, the invention features a method of treating a subject infected with a virus or at risk for or afflicted with a disorder or disease associated with a viral infection.

15 The method includes:

providing an siRNA, *e.g.*, and siRNA having a structure described herein, which siRNA is homologous to and can silence, *e.g.*, by cleavage, a viral gene or a cellular gene which mediates viral function, *e.g.*, entry or growth; and

20 administering the siRNA to a subject, preferably a human subject, thereby treating the subject.

Thus, the invention provides for a method of treating patients infected by the Human Papilloma Virus (HPV) or at risk for or afflicted with a disorder mediated by HPV, *e.g.*, cervical cancer. HPV is linked to 95% of cervical carcinomas and thus an antiviral therapy is an attractive method to treat these cancers and other symptoms of viral infection.

25 In a preferred embodiment, the expression of a HPV gene is reduced. In another preferred embodiment, the HPV gene is one of the group of E2, E6, or E7.

In a preferred embodiment, the expression of a human gene that is required for HPV replication is reduced.

30 The invention also includes a method of treating patients infected by the Human Immunodeficiency Virus (HIV) or at risk for or afflicted with a disorder mediated by HIV, *e.g.*, Acquired Immune Deficiency Syndrome (AIDS).

In a preferred embodiment, the expression of a HIV gene is reduced. In another preferred embodiment, the HIV gene is CCR5, Gag, or Rev.

In a preferred embodiment, the expression of a human gene that is required for HIV replication is reduced. In another preferred embodiment, the gene is CD4 or Tsg101.

5 The invention also includes a method for treating patients infected by the Hepatitis B Virus (HBV) or at risk for or afflicted with a disorder mediated by HBV, *e.g.*, cirrhosis and hepatocellular carcinoma.

In a preferred embodiment, the expression of a HBV gene is reduced. In another preferred embodiment, the targeted HBV gene encodes one of the group of the tail region of the HBV core protein, the pre-core region (pre-c) region, or the core region (c) region. In another  
10 preferred embodiment, a targeted HBV-RNA sequence is comprised of the poly(A) tail.

In a preferred embodiment, the expression of a human gene that is required for HBV replication is reduced.

The invention also provides for a method of treating patients infected by the Hepatitis  
15 A Virus (HAV), or at risk for or afflicted with a disorder mediated by HAV.

In a preferred embodiment, the expression of a human gene that is required for HAV replication is reduced.

The present invention provides for a method of treating patients infected by the Hepatitis C Virus (HCV), or at risk for or afflicted with a disorder mediated by HCV, *e.g.*,  
20 cirrhosis.

In a preferred embodiment, the expression of a HCV gene is reduced.

In another preferred embodiment the expression of a human gene required for HCV replication is reduced.

The present invention also provides for a method of treating patients infected by any  
25 of the group of Hepatitis viral strains comprising Hepatitis D, E, F, G, or H, or patients at risk for or afflicted with a disorder mediated by any of these strains of Hepatitis.

In a preferred embodiment, the expression of a Hepatitis D, E, F, G, or H gene is reduced.

In another preferred embodiment the expression of a human gene that is required for  
30 Hepatitis D, E, F, G or H replication is reduced.

Methods of the invention also provide for treating patients infected by the Respiratory Syncytial Virus (RSV) or at risk for or afflicted with a disorder mediated by RSV, *e.g.*, lower respiratory tract infection in infants and childhood asthma, pneumonia and other complications, *e.g.*, in the elderly.

5 In a preferred embodiment, the expression of a RSV gene is reduced. In another preferred embodiment, the targeted RSV gene encodes one of the group of genes N, L, or P.

In a preferred embodiment, the expression of a human gene that is required for RSV replication is reduced.

10 Methods of the invention provide for treating patients infected by the Herpes Simplex Virus (HSV) or at risk for or afflicted with a disorder mediated by HSV, *e.g.*, genital herpes and cold sores as well as life-threatening or sight-impairing diseases mainly in immunocompromised patients.

In a preferred embodiment, the expression of a HSV gene is reduced. In another preferred embodiment, the targeted HSV gene encodes DNA polymerase or the helicase-primase.

15 In a preferred embodiment, the expression of a human gene that is required for HSV replication is reduced.

The invention also provides a method for treating patients infected by the herpes Cytomegalovirus (CMV) or at risk for or afflicted with a disorder mediated by CMV, *e.g.*, congenital virus infections and morbidity in immunocompromised patients.

In a preferred embodiment, the expression of a CMV gene is reduced.

In a preferred embodiment, the expression of a human gene that is required for CMV replication is reduced.

25 Methods of the invention also provide for a method of treating patients infected by the herpes Epstein Barr Virus (EBV) or at risk for or afflicted with a disorder mediated by EBV, *e.g.*, NK/T-cell lymphoma, non-Hodgkin's lymphoma, and Hodgkin's disease.

In a preferred embodiment, the expression of a EBV gene is reduced.

In a preferred embodiment, the expression of a human gene required for EBV replication is reduced.

30 Methods of the invention also provide for treating patients infected by Kaposi's Sarcoma-associated Herpes Virus (KSHV), also called human herpesvirus 8, or patients at

risk for or afflicted with a disorder mediated by KSHV, e.g., Kaposi's sarcoma, multicentric Castleman's disease and AIDS-associated primary effusion lymphoma.

In a preferred embodiment, the expression of a KSHV gene is reduced.

In a preferred embodiment, the expression of a human gene that is required for KSHV replication is reduced.

The invention also includes a method for treating patients infected by the JC Virus (JCV) or a disease or disorder associated with this virus, e.g., progressive multifocal leukoencephalopathy (PML).

In a preferred embodiment, the expression of a JCV gene is reduced.

In preferred embodiment the expression of a human gene that is required for JCV replication is reduced.

Methods of the invention also provide for treating patients infected by the myxovirus or at risk for or afflicted with a disorder mediated by myxovirus; e.g., influenza.

In a preferred embodiment, the expression of a myxovirus gene is reduced.

In a preferred embodiment, the expression of a human gene that is required for myxovirus replication is reduced.

Methods of the invention also provide for treating patients infected by the rhinovirus or at risk for or afflicted with a disorder mediated by rhinovirus, e.g., the common cold.

In a preferred embodiment, the expression of a rhinovirus gene is reduced.

In preferred embodiment the expression of a human gene that is required for rhinovirus replication is reduced.

Methods of the invention also provide for treating patients infected by the coronavirus or at risk for or afflicted with a disorder mediated by coronavirus, e.g., the common cold.

In a preferred embodiment, the expression of a coronavirus gene is reduced.

In preferred embodiment the expression of a human gene that is required for coronavirus replication is reduced.

Methods of the invention also provide for treating patients infected by the flavivirus West Nile or at risk for or afflicted with a disorder mediated by West Nile Virus.

In a preferred embodiment, the expression of a West Nile Virus gene is reduced.

In another preferred embodiment, the West Nile Virus gene is one of the group comprising E, NS3, or NS5.

In a preferred embodiment, the expression of a human gene that is required for West Nile Virus replication is reduced.

Methods of the invention also provide for treating patients infected by the St. Louis Encephalitis flavivirus, or at risk for or afflicted with a disease or disorder associated with this virus, *e.g.*, viral haemorrhagic fever or neurological disease.

In a preferred embodiment, the expression of a St. Louis Encephalitis gene is reduced.

In a preferred embodiment, the expression of a human gene that is required for St. Louis Encephalitis virus replication is reduced.

Methods of the invention also provide for treating patients infected by the Tick-borne encephalitis flavivirus, or at risk for or afflicted with a disorder mediated by Tick-borne encephalitis virus, *e.g.*, viral haemorrhagic fever and neurological disease.

In a preferred embodiment, the expression of a Tick-borne encephalitis virus gene is reduced.

In a preferred embodiment, the expression of a human gene that is required for Tick-borne encephalitis virus replication is reduced.

Methods of the invention also provide for methods of treating patients infected by the Murray Valley encephalitis flavivirus, which commonly results in viral haemorrhagic fever and neurological disease.

In a preferred embodiment, the expression of a Murray Valley encephalitis virus gene is reduced.

In a preferred embodiment, the expression of a human gene that is required for Murray Valley encephalitis virus replication is reduced.

The invention also includes methods for treating patients infected by the dengue flavivirus, or a disease or disorder associated with this virus, *e.g.*, dengue haemorrhagic fever.

In a preferred embodiment, the expression of a dengue virus gene is reduced.

In a preferred embodiment, the expression of a human gene that is required for dengue virus replication is reduced.

Methods of the invention also provide for treating patients infected by the Simian Virus 40 (SV40) or at risk for or afflicted with a disorder mediated by SV40, *e.g.*, tumorigenesis.

In a preferred embodiment, the expression of a SV40 gene is reduced.

In a preferred embodiment, the expression of a human gene that is required for SV40 replication is reduced.

5 The invention also includes methods for treating patients infected by the Human T Cell Lymphotropic Virus (HTLV), or a disease or disorder associated with this virus, *e.g.*, leukemia and myelopathy.

In a preferred embodiment, the expression of a HTLV gene is reduced. In another preferred embodiment the HTLV gene is the Tax transcriptional activator.

10 In a preferred embodiment, the expression of a human gene that is required for HTLV replication is reduced.

Methods of the invention also provide for treating patients infected by the Moloney-Murine Leukemia Virus (Mo-MuLV) or at risk for or afflicted with a disorder mediated by Mo-MuLV, *e.g.*, T-cell leukemia.

In a preferred embodiment, the expression of a Mo-MuLV gene is reduced.

15 In a preferred embodiment, the expression of a human gene that is required for Mo-MuLV replication is reduced.

Methods of the invention also provide for treating patients infected by the encephalomyocarditis virus (EMCV) or at risk for or afflicted with a disorder mediated by EMCV, *e.g.*, myocarditis. EMCV leads to myocarditis in mice and pigs and is capable of  
20 infecting human myocardial cells. This virus is therefore a concern for patients undergoing xenotransplantation.

In a preferred embodiment, the expression of a EMCV gene is reduced.

In a preferred embodiment, the expression of a human gene that is required for EMCV replication is reduced.

25 The invention also includes a method for treating patients infected by the measles virus (MV) or at risk for or afflicted with a disorder mediated by MV, *e.g.*, measles.

In a preferred embodiment, the expression of a MV gene is reduced.

In a preferred embodiment, the expression of a human gene that is required for MV replication is reduced.



The invention also includes a method for treating patients infected by the Varicella zoster virus (VZV) or at risk for or afflicted with a disorder mediated by VZV, *e.g.*, chicken pox or shingles (also called zoster).

In a preferred embodiment, the expression of a VZV gene is reduced.

5 In a preferred embodiment, the expression of a human gene that is required for VZV replication is reduced.

The invention also includes a method for treating patients infected by an adenovirus or at risk for or afflicted with a disorder mediated by an adenovirus, *e.g.*, respiratory tract infection.

10 In a preferred embodiment, the expression of an adenovirus gene is reduced.

In a preferred embodiment, the expression of a human gene that is required for adenovirus replication is reduced.

The invention includes a method for treating patients infected by a yellow fever virus (YFV) or at risk for or afflicted with a disorder mediated by a YFV, *e.g.*, respiratory tract infection.

15 In a preferred embodiment, the expression of a YFV gene is reduced. In another preferred embodiment, the preferred gene is one of a group that includes the E, NS2A, or NS3 genes.

In a preferred embodiment, the expression of a human gene that is required for YFV replication is reduced.

20 Methods of the invention also provide for treating patients infected by the poliovirus or at risk for or afflicted with a disorder mediated by poliovirus, *e.g.*, polio.

In a preferred embodiment, the expression of a poliovirus gene is reduced.

In a preferred embodiment, the expression of a human gene that is required for poliovirus replication is reduced.

25 Methods of the invention also provide for treating patients infected by a poxvirus or at risk for or afflicted with a disorder mediated by a poxvirus, *e.g.*, smallpox.

In a preferred embodiment, the expression of a poxvirus gene is reduced.

In a preferred embodiment, the expression of a human gene that is required for poxvirus replication is reduced.

In another aspect, the invention features methods of treating a subject infected with a pathogen, *e.g.*, a bacterial, amoebic, parasitic, or fungal pathogen. The method includes:

providing an siRNA, *e.g.*, an siRNA having a structure described herein, where siRNA is homologous to and can silence, *e.g.*, by cleavage of a pathogen gene; and  
5 administering the siRNA to a subject, preferably a human subject, thereby treating the subject.

The target gene can be one involved in growth, cell wall synthesis, protein synthesis, transcription, energy metabolism (*e.g.*, the Krebs cycle), or toxin production. Thus, the present invention provides for a method of treating patients infected by a plasmodium that  
10 causes malaria.

In a preferred embodiment, the expression of a plasmodium gene is reduced. In another preferred embodiment, the gene is apical membrane antigen 1 (AMA1).

In a preferred embodiment, the expression of a human gene that is required for plasmodium replication is reduced.

15 The invention also includes methods for treating patients infected by the *Mycobacterium ulcerans*, or a disease or disorder associated with this pathogen, *e.g.*, a Buruli ulcer.

In a preferred embodiment, the expression of a *Mycobacterium ulcerans* gene is reduced.

20 In a preferred embodiment, the expression of a human gene that is required for *Mycobacterium ulcerans* replication is reduced.

The invention also includes methods for treating patients infected by *Mycobacterium tuberculosis*, or a disease or disorder associated with this pathogen, *e.g.*, tuberculosis.

25 In a preferred embodiment, the expression of a *Mycobacterium tuberculosis* gene is reduced.

In a preferred embodiment, the expression of a human gene that is required for *Mycobacterium tuberculosis* replication is reduced.

The invention also includes methods for treating patients infected by *Mycobacterium leprae*, or a disease or disorder associated with this pathogen, *e.g.*, leprosy.

30 In a preferred embodiment, the expression of a *Mycobacterium leprae* gene is reduced.

In a preferred embodiment, the expression of a human gene that is required for *Mycobacterium leprae* replication is reduced.

The invention also includes methods for treating patients infected by the bacteria *Staphylococcus aureus*, or a disease or disorder associated with this pathogen, *e.g.*, infections  
5 of the skin or mucous membranes.

In a preferred embodiment, the expression of a *Staphylococcus aureus* gene is reduced.

In a preferred embodiment, the expression of a human gene that is required for *Staphylococcus aureus* replication is reduced.

10 The invention also includes methods for treating patients infected by the bacteria *Streptococcus pneumoniae*, or a disease or disorder associated with this pathogen, *e.g.*, pneumonia or childhood lower respiratory tract infection.

In a preferred embodiment, the expression of a *Streptococcus pneumoniae* gene is reduced.

15 In a preferred embodiment, the expression of a human gene that is required for *Streptococcus pneumoniae* replication is reduced.

The invention also includes methods for treating patients infected by the bacteria *Streptococcus pyogenes*, or a disease or disorder associated with this pathogen, *e.g.*, Strep throat or Scarlet fever.

20 In a preferred embodiment, the expression of a *Streptococcus pyogenes* gene is reduced.

In a preferred embodiment, the expression of a human gene that is required for *Streptococcus pyogenes* replication is reduced.

25 The invention also includes methods for treating patients infected by the bacteria *Chlamydia pneumoniae*, or a disease or disorder associated with this pathogen, *e.g.*, pneumonia or childhood lower respiratory tract infection.

In a preferred embodiment, the expression of a *Chlamydia pneumoniae* gene is reduced.

30 In a preferred embodiment, the expression of a human gene that is required for *Chlamydia pneumoniae* replication is reduced.

The invention also includes methods for treating patients infected by the bacteria *Mycoplasma pneumoniae*, or, a disease or disorder associated with this pathogen, *e.g.*, pneumonia or childhood lower respiratory tract infection.

5 In a preferred embodiment, the expression of a *Mycoplasma pneumoniae* gene is reduced.

In a preferred embodiment, the expression of a human gene that is required for *Mycoplasma pneumoniae* replication is reduced.

In one aspect, the invention features, a method of treating a subject, *e.g.*, a human, at risk for or afflicted with a disease or disorder characterized by an unwanted immune response, *e.g.*, an inflammatory disease or disorder, or an autoimmune disease or disorder.  
10 The method includes:

providing an siRNA, *e.g.*, an siRNA having a structure described herein, which siRNA is homologous to and can silence, *e.g.*, by cleavage, a gene which mediates an unwanted immune response; and

15 administering the siRNA to a subject, thereby treating the subject.

In a preferred embodiment, the disease or disorder is an ischemia or reperfusion injury, *e.g.*, ischemia reperfusion or injury associated with acute myocardial infarction, unstable angina, cardiopulmonary bypass, surgical intervention (*e.g.*, angioplasty, such as percutaneous transluminal coronary angioplasty), a response to a transplanted organ or tissue  
20 (*e.g.*, transplanted cardiac or vascular tissue), or thrombolysis.

In a preferred embodiment, the disease or disorder is restenosis, *e.g.*, restenosis associated with surgical intervention (*e.g.*, angioplasty, such as percutaneous transluminal coronary angioplasty).

In a preferred embodiment, the disease or disorder is Inflammatory Bowel Disease, *e.g.*, Crohn's Disease or Ulcerative Colitis.  
25

In a preferred embodiment, the disease or disorder is inflammation associated with an infection or injury.

In a preferred embodiment, the disease or disorder is asthma, lupus, multiple sclerosis, diabetes, *e.g.*, type II diabetes, arthritis, *e.g.*, rheumatoid or psoriatic.

30 In particularly preferred embodiments, the siRNA silences an integrin or co-ligand thereof, *e.g.*, VLA4, VCAM, ICAM.

In particularly preferred embodiments, the siRNA silences a selectin or co-ligand thereof, *e.g.*, P-selectin, E-selectin (ELAM), I-selectin, or P-selectin glycoprotein (PSGL1).

In particularly preferred embodiments, the siRNA silences a component of the complement system, *e.g.*, C3, C5, C3aR, C5aR, C3 convertase, C5 convertase.

5 In particularly preferred embodiments, the siRNA silences a chemokine or receptor thereof, *e.g.*, TNF $\alpha$ , TNF $\beta$ , IL-1 $\alpha$ , IL-1 $\beta$ , IL-2, IL-2R, IL-4, IL-4R, IL-5, IL-6, IL-8, TNFRI, TNFRII, IgE, SCYA11, or CCR3.

In less preferred embodiments the siRNA silences GCSF, Gro1, Gro2, Gro3, PF4, MIG, Pro-Platelet Basic Protein (PPBP), MIP-1 $\alpha$ , MIP-1 $\beta$ , RANTES, MCP-1, MCP-2,  
10 MCP-3, CMBKR1, CMBKR2, CMBKR3, CMBKR5, AIF-1, or I-309.

In one aspect, the invention features, a method of treating a subject, *e.g.*, a human, at risk for or afflicted with acute pain or chronic pain. The method includes:

providing an siRNA, *e.g.*, an siRNA having a structure described herein, which siRNA is homologous to and can silence, *e.g.*, by cleavage, a gene which mediates the  
15 processing of pain; and

administering the siRNA to a subject, thereby treating the subject.

In particularly preferred embodiments, the siRNA silences a component of an ion channel.

In particularly preferred embodiments, the siRNA silences a neurotransmitter receptor  
20 or ligand.

In one aspect, the invention features a method of treating a subject, *e.g.*, a human, at risk for or afflicted with a neurological disease or disorder. The method includes:

providing an siRNA, *e.g.*, an siRNA having a structure described herein, which siRNA is homologous to and can silence, *e.g.*, by cleavage, a gene which mediates a  
25 neurological disease or disorder; and

administering the siRNA to a subject, thereby treating the subject.

In a preferred embodiment, the disease or disorder is Alzheimer's Disease or Parkinson's Disease.

In particularly preferred embodiments, the siRNA silences an amyloid-family gene,  
30 *e.g.*, APP; a presenilin gene, *e.g.*, PSEN1 and PSEN2, or  $\alpha$ -synuclein.

In a preferred embodiment, the disease or disorder is a neurodegenerative trinucleotide repeat disorder, *e.g.*, Huntington's disease, dentatorubral pallidolysian atrophy or a spinocerebellar ataxia, *e.g.*, SCA1, SCA2, SCA3 (Machado-Joseph disease), SCA7 or SCA8.

5 In particularly preferred embodiments, the siRNA silences HD, DRPLA, SCA1, SCA2, MJD1, CACNLIA4, SCA7, or SCA8.

The loss of heterozygosity (LOH) can result in hemizyosity for sequence, *e.g.*, genes, in the area of LOH. This can result in a significant genetic difference between normal and disease-state cells, *e.g.*, cancer cells, and provides a useful difference between normal and disease-state cells, *e.g.*, cancer cells. This difference can arise because a gene or other  
10 sequence is heterozygous in euploid cells but is hemizygous in cells having LOH.

The regions of LOH will often include a gene, the loss of which promotes unwanted proliferation, *e.g.*, a tumor suppressor gene, and other sequences, such as other genes and, in some cases, a gene which is essential for normal function, *e.g.*, growth.

15 Methods of the invention rely, in part, on the specific cleavage or silencing of one allele of an essential gene with an siRNA of the invention. The siRNA is selected such that it targets the single allele of the essential gene found in the cells having LOH but does not silence the other allele, which is present in cells which do not show LOH. In essence, it discriminates between the two alleles, preferentially silencing the selected allele. In essence, polymorphisms, *e.g.*, SNPs, of essential genes that are affected by LOH, are used targets for  
20 a disorder characterized by cells having LOH, *e.g.*, cancer cells having LOH.

*E.g.*, one of ordinary skill in the art can identify essential genes which are in proximity to tumor suppressor genes, and which are within a LOH region which includes the tumor suppressor gene. The gene encoding the large subunit of human RNA polymerase II, POLR2A, a gene located in close proximity to the tumor suppressor gene p53, is such a gene.  
25 It frequently occurs within a region of LOH in cancer cells. Other genes that occur within LOH regions, and are lost in many cancer cell types, include the group comprising replication protein A 70-kD subunit, replication protein A 32-kD, ribonucleotide reductase, thymidilate synthase, TATA associated factor 2H, ribosomal protein S14, eukaryotic initiation factor 5A, alanyl tRNA synthetase, cysteinyl tRNA synthetase, NaK ATPase,  
30 alpha-1 subunit, and transferrin receptor.

Accordingly, the invention features a method of treating a disorder characterized by LOH, *e.g.*, cancer. The method includes, optionally, determining the genotype of the allele of a gene in the region of LOH and preferably determining the genotype of both alleles of the gene in a normal cell; providing an siRNA which preferentially cleaves or silences the allele found in the LOH cells; and administering the siRNA to the subject, thereby treating the disorder.

The invention also includes an siRNA disclosed herein, *e.g.*, an siRNA which can preferentially silence, *e.g.*, cleave, one allele of a polymorphic gene.

In another aspect, the invention provides a method of cleaving or silencing more than one gene with an siRNA. In these embodiments the siRNA is selected so that it has sufficient homology to a sequence found in more than one gene. For example, the sequence AAGCTGGCCCTGGACATGGAGAT is conserved between mouse lamin B1, lamin B2, keratin complex 2-gene 1 and lamin A/C. Thus an siRNA targeted to this sequence would effectively silence this entire collection of genes.

The invention also includes an siRNA disclosed herein, *e.g.*, an siRNA which can silence more than one gene.

In another aspect, the invention includes vectors, *e.g.*, expression vectors, which encode siRNAs, *e.g.*, siRNAs homologous with a gene disclosed herein. These include vectors which can express one or both strands of one or more siRNAs; a composition which includes a first vector which encodes a first strand of an siRNA and a second vector which encodes a second strand of an siRNA; a vector which encodes a hairpin siRNA, *e.g.*, a hairpin which upon cleavage provides both strands of an siRNA. The vector can also encode, and preferably express, a protein, *e.g.*, a protein active in siRNA metabolism or function, *e.g.*, Dicer or Ago 2.

In a preferred embodiment, the vector or delivery method, *e.g.*, a virus, is selected such that the siRNA integrates into a preselected site, *e.g.*, a site which will result in modulated, *e.g.*, inducible or controlled expression, *e.g.*, by a pol III promoter or by a pol II promoter containing temporal, developmental, disease-state, or tissue specific promoter elements. In other embodiments, the expression of an siRNA strand is driven by a temporal, developmental, disease-state, or tissue specific promoter, *e.g.*, a pol II promoter containing temporal, developmental, disease-state or tissue specific elements.

Unless otherwise defined, all technical and scientific terms used herein have the same meaning as commonly understood by one of ordinary skill in the art to which this invention belongs. Although methods and materials similar or equivalent to those described herein can be used in the practice or testing of the present invention, useful methods and materials are described below. The materials, methods, and examples are illustrative only and not intended to be limiting. Other features and advantages of the invention will be apparent from the following detailed description and from the claims.

### BRIEF DESCRIPTION OF THE DRAWINGS

**FIG. 1A** is a graphical representation of dsRNAs used for targeting Pp-luc mRNA. Three series of blunt-ended dsRNAs covering a range of 29 to 504 bp were prepared. The position of the first nucleotide of the sense strand of the dsRNA is indicated relative to the start codon of Pp-luc mRNA (pl).

**FIG. 1B** is a graphical representation of the ratios of target Pp-luc to control Rr-luc activity when normalized to a buffer control (buf). DsRNAs (5 nM) were preincubated in *Drosophila* lysate for 15 min at 25°C prior to the addition of 7-methyl-guanosine-capped Pp-luc and Rr-luc mRNA (~50 pM). The incubation was continued for another hour and then analyzed by the dual luciferase assay (Promega). The data are the average from at least four independent experiments standard deviation.

**FIG. 2** is a gel from a time course experiment of 21-23mer formation in the processing of internally <sup>32</sup>P-labeled dsRNAs (5 nM) in *Drosophila* lysate. The length and source of the dsRNA are indicated. An RNA size marker (M) was loaded in the left lane and the fragment sizes are indicated. Double bands at time zero are due to incompletely denatured dsRNA.

**FIG. 3A** is a denaturing gel of the stable 5' cleavage products produced by 1 h incubation of 10 nM sense or antisense RNA <sup>32</sup>P-labeled at the cap with 10 nM dsRNAs of the p133 series in *Drosophila* lysate. Length markers were generated by partial nuclease T1 digestion and partial alkaline hydrolysis (OH) of the cap-labeled target RNA. The regions targeted by the dsRNAs are indicated as vertical bars on the left and right sides of the gel. The 20-23 nt spacing between the predominant cleavage sites for the 111 bp long dsRNA is shown. The horizontal arrow indicates unspecific cleavage not due to RNAi.



FIG. 3B shows the sequences of the capped 177 nt sense and 180 nt antisense target RNAs, in antiparallel orientation, such that complementary sequences are opposing each other. The regions targeted by the different dsRNAs are indicated by black bars positioned between the sense and antisense target sequences. Cleavage sites are indicated by circles: large circles for strong cleavage, and small circles for weak cleavage. The  $^{32}\text{P}$ -radiolabeled phosphate group is marked by an asterisk.

FIG. 4A shows the sequences of ~21 nt RNAs after dsRNA processing. The ~21 nt RNA fragments were directionally cloned and sequenced. Some oligoribonucleotides originated from the sense strand of the dsRNA, while others originated from the antisense strand. Thick bars represent sequences present in multiple clones, and the number at the right indicates the frequency. The target RNA cleavage sites mediated by the dsRNA are indicated as circles (large circles for strong cleavage, small circles for weak cleavage (see FIG. 3B)). Circles on top of the sense strand indicated cleavage sites within the sense target and circles at the bottom of the dsRNA indicate cleavage site in the antisense target. Up to five additional nucleotides were identified in ~21 nt fragments derived from the 3' ends of the dsRNA. These nucleotides are random combinations of predominantly C, G, or A residues.

FIG 4B is a collage of two-dimensional TLC plates. ~21 nt RNAs were generated by incubation of internally radiolabeled 504 bp Pp-luc dsRNA in *Drosophila* lysate, gel-purified, and then digested to mononucleotides with nuclease P1 (top row) or ribonuclease T2 (bottom row). The dsRNA was internally radiolabeled by transcription in the presence of one of the indicated  $\alpha$ - $^{32}\text{P}$  nucleoside triphosphates. Radioactivity was detected by phosphorimaging. Nucleoside 5'-monophosphates, nucleoside 3'-monophosphates, nucleoside 5', 3'-diphosphates, and inorganic phosphates are indicated as pN, Np, pNp, and pi, respectively. Hollow circles indicate UV-absorbing spots from non-radioactive carrier nucleotides. The 3', 5' bisphosphates were identified by co-migration with radiolabeled standards prepared by 5'-phosphorylation of nucleoside 3'-monophosphates with T4 polynucleotide kinase and  $\gamma$ - $^{32}\text{P}$ -ATP.

FIG. 5A depicts a control 52 bp dsRNA and synthetic 21 and 22 nt dsRNAs. The sequences of the siRNAs were derived from the cloned fragments of 52 and 111 bp dsRNAs (FIG. 4A), except for the 22 nt antisense strand of duplex 5. The siRNAs in duplexes 6 and 7 were unique to the 111 bp dsRNA processing reaction. Both strands of the control 52 bp

dsRNA were prepared by *in vitro* transcription and a fraction of transcripts may contain untemplated 3' nucleotide addition. The target RNA cleavage sites directed by the siRNA duplexes are indicated by circles (see legend to FIG. 4A) and were determined as described in FIG. 5B.

5        **FIG. 5B** is a gel showing the position of cleavage sites on the sense and antisense target RNAs. The target RNA sequences are as described in FIG. 3B. A control 52 bp dsRNA (10 nM), or 21 and 22 nt RNA duplexes 1-7 (100 nM) (see FIG. 5A) were incubated with target RNA for 2.5 h at 25°C in *Drosophila* lysate. The stable 5' cleavage products were resolved on the gel. The cleavage sites are indicated in FIG. 5A. The region targeted by the  
10        52 bp dsRNA or the sense (s) or antisense (as) strands are indicated by the vertical bars on the right and left side of the gel. The cleavage sites are all located within the region of identity of the dsRNAs.

**FIG. 6A** is a representation of 52 bp dsRNA constructs. The overhang regions are 3' extensions of the sense and antisense strands. The observed cleavage sites on the target  
15        RNAs are represented as circles as in FIG. 4A and were determined as shown in FIG. 6B.

**FIG. 6B** is a gel showing the position of the cleavage sites on the sense and antisense target RNAs. The target RNA sequences are as described in FIG. 3B. DsRNA (10 nM) was incubated with target RNA for 2.5 h at 25°C in *Drosophila* lysate. The stable 5' cleavage products were resolved on the gel. The major cleavage sites are indicated with a horizontal  
20        arrow and are also represented in FIG. 6A. The region targeted by the 52 bp dsRNA is represented as a vertical bar on either side of the gel.

**FIG. 7** is a model for RNAi. According to this model, RNAi begins with processing of dsRNA to predominantly 21 and 22 nt short interfering RNAs (siRNAs). Short overhanging 3' nucleotides associate with the dsRNA and facilitate processing of the short  
25        dsRNAs. The dsRNA-processing proteins, which remain to be characterized, are represented as ovals, and are shown assembled on the dsRNA in asymmetric fashion. For example, one protein or protein domain may always associate with the siRNA strand in the 3' to 5' direction, or on the antisense strand while another protein may always associate with the opposing, or sense, siRNA strand. These proteins or a subset of the proteins may remain  
30        associated with the siRNA duplex and preserve its orientation as determined by the direction of the dsRNA processing reaction. Only the siRNA sequence associated with the one

particular protein may be able to guide target RNA cleavage. The endonuclease complex is referred to as small interfering ribonucleoprotein complex, or siRNP. It is presumed here that the endonuclease that cleaves the dsRNA may also cleave the target RNA, probably by temporarily displacing the passive siRNA strand, which is not used for target recognition.

5 The target RNA can then be cleaved in the center of the region recognized by the sequence-complementary guide siRNA.

FIG. 8A is a representation of the firefly (Pp-luc) and sea pansy (Rr-luc) luciferase reporter gene regions from plasmids pGL2-Control, pGL3-Control and pRL-TK (Promega). SV40 regulatory elements, the HSV thymidine kinase promoter and the two introns (lines)

10 are indicated. The sequence of GL3 luciferase is 95% identical to GL2, but RL is completely unrelated to both. Luciferase expression from pGL2 is approx. 10-fold lower than from pGL3 in transfected mammalian cells. The region targeted by the siRNA duplexes is indicated as a black bar below the coding region of the luciferase genes.

FIG. 8B illustrates the sequences of the sense (top) and antisense (bottom) strands of

15 the siRNA duplexes targeting GL2, GL3 and RL luciferase RNAs. The GL2 and GL3 siRNA duplexes differ by only 3 single nucleotide substitutions (boxed). As an unspecific control, a duplex with the inverted GL2 sequence, invGL2, was synthesized. The 2 nt 3' overhang of 2'-deoxythymidine is indicated as TT; uGL2 is similar to GL2 siRNA but contains ribo-uridine 3' overhangs.

FIG. 9 is a collage of graphs indicating ratios of target control luciferase normalized to a buffer control (bu); gray bars indicate ratios of *Photinus pyralis* (Pp-luc) GL2 or GL3 luciferase to *Renilla reniformis* (Rr-luc) RL luciferase (left axis), white bars indicate RL to GL2 or GL3 ratios (right axis). Panels a, c, e, g, and i describe experiments performed with the combination of pGL2-Control and pRL-TK reporter plasmids, and panels b, d, f, h and j

25 describe experiments performed with pGL3-Control and pRL-TK reporter plasmids. The cell line used for the interference experiment is indicated at the top of each plot. The ratios of Pp-luc/Rr-luc for the buffer control (bu) varied between 0.5 and 10 for pGL2/pRL and between 0.03 and 1 for pGL3/pRL, respectively, before normalization and between the various cell lines tested. The plotted data were averaged from three independent experiments

30 S.D.

**FIG. 10** is a collage of graphs illustrating the lengths of long dsRNAs (x axis). Panels a, c and e describe experiments performed with pGL2-Control and pRL-TK reporter plasmids, panels b, d and f with pGL3-Control and pRL-TK reporter plasmids. The data were averaged from two independent experiments S.D. (a), (b) Absolute Pp-luc expression, plotted in arbitrary luminescence units. (c), (d) Rr-luc expression, plotted in arbitrary luminescence units. (e), (f) Ratios of normalized target to control luciferase. The ratios of luciferase activity for siRNA duplexes were normalized to a buffer control (bu). The luminescence ratios for 50 or 500 bp dsRNAs were normalized to the respective ratios observed for 50 and 500 bp dsRNA from humanized GFP (hG). The overall differences in sequences between the 49 and 484 bp dsRNAs targeting GL2 and GL3 were not sufficient to confer specificity between GL2 and GL3 targets (43 nt uninterrupted identity in 49 bp segment, 239 nt longest uninterrupted identity in 484 bp segment).

**FIG. 11A** is an outline of the experimental strategy. The capped and polyadenylated sense target mRNA is depicted and the relative positions of sense and antisense siRNAs are shown. Eight series of duplexes, according to the eight different antisense strands were prepared. The siRNA sequences and the number of overhanging nucleotides were changed in 1-nt steps.

**FIG. 11B** is a graph illustrating the normalized relative luminescence of target luciferase (*Photinus pyralis*, Pp-luc) to control luciferase (*Renilla reniformis*, Rr-luc) in *D. melanogaster* embryo lysate in the presence of 5 nM blunt-ended dsRNAs. The luminescence ratios determined in the presence of dsRNA were normalized to the ratio obtained for a buffer control (bu). Normalized ratios less than 1 indicate specific interference.

**FIG. 11C (c-j)** is a collage of graphs illustrating the normalized interference ratios for eight series of 21-nt siRNA duplexes. The sequences of siRNA duplexes are depicted above the bar graphs. Each panel shows the interference ratio for a set of duplexes formed with a given antisense guide siRNA and 5 different sense siRNAs. The number of overhanging nucleotides (3' overhang, positive numbers; 5' overhangs, negative numbers) is indicated on the x-axis. Data points were averaged from at least 3 independent experiments, error bars represent standard deviations.

FIG. 12 is a diagram illustrating an experiment, and three graphs illustrating the results. Three 21-nt antisense strands were paired with eight sense siRNAs. The siRNAs were changed in length at their 3' end. The 3' overhang of the antisense siRNA was 1-nt (B), 2-nt (C), or 3-nt (D) while the sense siRNA overhang was varied for each series. The sequences of the siRNA duplexes and the corresponding interference ratios are indicated.

FIG. 13 is a diagram illustrating an experiment, and three graphs illustrating the results. The 21-nt siRNA duplex is identical in sequence to the one shown in FIG. 11H or 12C. The siRNA duplexes were extended to the 3' side of the sense siRNA (B) or the 5' side of the sense siRNA (C). The siRNA duplex sequences and the respective interference ratios are indicated.

FIG. 14 is a graph of luminescence ratios. The 2'-hydroxyl groups (OH) in the strands of siRNA duplexes were replaced by 2' deoxy (d) or 2'-O-methyl (Me). 2-nt and 4-nt 2'-deoxysubstitutions at the 3'-ends are indicated as 2-nt d and 4-nt d, respectively. Uridine residues were replaced by 2'-deoxy thymidine.

FIG. 15A is a diagram of <sup>32</sup>P (asterisk) cap-labeled sense and antisense target RNAs and siRNA duplexes. The positions of sense and antisense target RNA cleavage is indicated by triangles over and below the siRNA duplexes, respectively.

FIG. 15B is a pair of gels indicating target RNA cleavage sites. After 2 h incubation of 10 nM target with 100 nM siRNA duplex in *D. melanogaster* embryo lysate, the 5' cap-labeled substrate and the 5' cleavage products were resolved on sequencing gels. Length markers were generated by partial RNase T1 digestion (T1) and partial alkaline hydrolysis (OH-) of the target RNAs. The bold lines to the left of the images indicate the region covered by the siRNA strands 1 and 5 of the same orientation as the target.

FIG. 16 is a diagrammatic representation of the experiment (A, B) and a collage of gels showing the results (C, D). The antisense siRNA was the same length in all siRNA duplexes, but the sense strand was varied between 18 to 25 nt by changing the 3' end (A) or 18 to 23 nt by changing the 5' end (B). The position of sense and antisense target RNA cleavage is indicated by triangles over and below the siRNA duplexes, respectively. Analysis of target RNA cleavage was performed by using cap-labeled sense (C, D: top panel) or antisense (C, D: bottom panel) target RNAs. Only the cap-labeled 5' cleavage products are shown. The sequences of the siRNA duplexes are indicated, and the length of the sense

siRNA strands is marked on over the gels. The control lane, marked with a dash in panel (C), shows target RNA incubated in the absence of siRNAs. Markers were as described in FIG. 15. The arrows in (D), bottom panel, indicate the target RNA cleavage sites that differed by 1 nt.

5        **FIG. 17** is a graph illustrating luminescence ratios when the 2-nt 3'overhang (NN) was changed in sequence and composition as indicated (T, 2'-deoxythymidine, dG, 2'deoxyguanosine; asterisk, wild-type siRNA duplex). Normalized interference ratios were determined as described in FIG. 11. The wild-type sequence is the same as depicted in FIG. 14.

10        **FIG. 18** is a graph illustrating luminescence ratios resulting from experiments with mismatched siRNA duplexes. The sequences of the mismatched siRNA duplexes are shown above the graph; modified sequence segments or single nucleotides are shaded. The reference duplex (ref) and the siRNA duplexes 1 to 7 contain 2'-deoxythymidine 2-nt overhangs. The silencing efficiency of the thymidine-modified reference duplex was  
15        comparable to the wild-type sequence (FIG. 17). Normalized interference ratios were determined as described in FIG. 11.

**FIG. 19** is a collage of graphs illustrating luminescence ratios when siRNA duplexes were extended to the 3' side of the sense siRNA (A) or the 5' side of the sense siRNA (B). The siRNA duplex sequences and the respective interference ratios are indicated. For HeLa  
20        SS6 cells, siRNA duplexes (0.84  $\mu$ g) targeting GL2 luciferase were transfected together with pGL2-Control and pRL-TK plasmids. For comparison, the *in vitro* RNAi activities of siRNA duplexes tested in *D. melanogaster* lysate are indicated.

#### DETAILED DESCRIPTION

25        The invention relates to sequence and structural features of double stranded (ds)RNA molecules required to mediate target-specific nucleic acid modifications such as RNA-interference and/or DNA methylation. Novel dsRNA agents and methods capable of mediating target-specific RNA interference (RNAi) or other target-specific nucleic acid modifications, such as DNA methylation, are provided, and these agents and methods have  
30        improved efficacy and safety compared to prior art agents and methods.

The invention provides an isolated double stranded RNA molecule, wherein each RNA strand has a length from 19-25, particularly from 19-23 nucleotides, wherein said RNA molecule is capable of mediating target-specific nucleic acid modifications, particularly RNA interference and/or DNA methylation. Preferably at least one strand has a 3'-overhang from 1-5 nucleotides, more preferably from 1-3 nucleotides and most preferably 2 nucleotides. The other strand may be blunt-ended or has up to 6 nucleotides 3' overhang. Also, if both strands of the dsRNA are exactly 21 or 22 nt, it is possible to observe some RNA interference when both ends are blunt (0 nt overhang). The RNA molecule is preferably a synthetic RNA molecule that is substantially free from contaminants occurring in cell extracts, *e.g.*, from *Drosophila* embryos. Further, the RNA molecule is can be substantially free from any non-target-specific contaminants, particularly non-target-specific RNA molecules, *e.g.*, from contaminants occurring in cell extracts.

The invention also relates to the use of isolated double stranded RNA molecules, wherein each RNA strand has a length from 19-25 nucleotides, for mediating target specific nucleic acid modifications, particularly RNAi, in mammalian cells, particularly in human cells.

It was found that synthetic short double stranded RNA molecules particularly with overhanging 3' ends are sequence-specific mediators of RNAi and mediate efficient target-RNA cleavage, wherein the cleavage site is located near the center of the region spanned by the guiding short RNA.

Each strand of the RNA molecule preferably has a length from 20-22 nucleotides (or 20-25 nucleotides in mammalian cells), wherein the length of each strand may be the same or different. The length of the 3'-overhang preferably reaches from 1-3 nucleotides, wherein the length of the overhang may be the same or different for each strand. The RNA-strands preferably have 3'-hydroxyl groups. The 5'-terminus preferably comprises a phosphate, diphosphate, triphosphate or hydroxyl group. The most effective dsRNAs are composed of two 21 nt strands which are paired such that 1-3, particularly 2 nt 3' overhangs are present on both ends of the dsRNA.

The target RNA cleavage reaction guided by siRNAs is highly sequence-specific. However, not all positions of an siRNA contribute equally to target recognition. Mismatches in the center of the siRNA duplex are most critical and essentially abolish target RNA

cleavage. In contrast, the 3' nucleotide of the siRNA strand (*e.g.*, position 21) that is complementary to the single-stranded target RNA, does not contribute to specificity of the target recognition. Further, the sequence of the unpaired 2-nt 3' overhang of siRNA the strand with the same polarity as the target RNA is not critical for target RNA cleavage as only the antisense siRNA strand guides target recognition. Thus, from the single-stranded overhanging nucleotides only the penultimate position of the antisense siRNA (*e.g.*, position 20) needs to match the targeted sense mRNA.

The double stranded RNA molecules of the present invention exhibit a high *in vivo* stability in serum or in growth medium for cell cultures. In order to further enhance the stability, the 3'-overhangs may be stabilized against degradation, *e.g.*, they may be selected such that they consist of purine nucleotides, particularly adenosine or guanosine nucleotides. Alternatively, substitution of pyrimidine nucleotides by modified analogues, *e.g.*, substitution of uridine 2 nt 3' overhangs by 2'-deoxythymidine is tolerated and does not affect the efficiency of RNA interference. The absence of a 2' hydroxyl significantly enhances the nuclease resistance of the overhang in tissue culture medium.

The RNA molecule of the invention preferably contains at least one modified nucleotide analogue. The nucleotide analogues may be located at positions where the target-specific activity, *e.g.*, the RNAi mediating activity is not substantially effected, *e.g.*, in a region at the 5'-end and/or the 3'-end of the double stranded RNA molecule. Particularly, the overhangs may be stabilized by incorporating modified nucleotide analogues.

Preferred nucleotide analogues are selected from sugar- or backbone-modified ribonucleotides. It should be noted, however, that also nucleobase-modified ribonucleotides, *i.e.*, ribonucleotides, containing a non-naturally occurring nucleobase instead of a naturally occurring nucleobase, such as uridines or cytidines modified at the 5-position, *e.g.*, 5-(2-amino)propyl uridine, 5-bromo uridine; adenosines and guanosines modified at the 8-position, *e.g.*, 8-bromo guanosine; deaza nucleotides, *e.g.*, 7-deazaadenosine; O- and N-alkylated nucleotides, *e.g.*, N6-methyl adenosine are suitable. In preferred sugar-modified ribonucleotides, the 2' OH-group can be replaced by a group selected from H, OR, R halo, SH, SR, NH<sub>2</sub>, NHR, NR<sub>2</sub> or CN, wherein R is C1-C6 alkyl, alkenyl or alkynyl and halo is F, Cl, Br or I.



In preferred backbone-modified ribonucleotides, the phosphoester group connecting to adjacent ribonucleotides is replaced by a modified group, *e.g.*, a phosphothioate group. It should be noted that the above modifications may be combined.

The sequence of the double stranded RNA molecule of the present invention has to  
5 have a sufficient identity to a nucleic acid target molecule in order to mediate target-specific RNAi and/or DNA methylation. Preferably, the sequence has an identity of at least 50%, particularly of at least 70% to the desired target molecule in the double stranded portion of the RNA molecule. More preferably, the identity is at least 85% and most preferably 100% in the double stranded portion of the RNA molecule. The identity of a double stranded RNA  
10 molecule to a predetermined nucleic acid target molecule, *e.g.*, an mRNA target molecule, may be determined as follows:

$$I = (n/L) \times 100$$

wherein I is the identity in percent, n is the number of identical nucleotides in the double stranded portion of the ds RNA and the target, and L is the length of the sequence overlap of  
15 the double stranded portion of the dsRNA and the target.

Alternatively, the identity of the double stranded RNA molecule to the target sequence may be defined including the 3' overhang, particularly an overhang having a length from 1-3 nucleotides. In this case the sequence identity is preferably at least 50%, more preferably at least 70% and most preferably at least 85% to the target sequence. For  
20 example, the nucleotides from the 3' overhang and up to 2 nucleotides from the 5' and/or 3' terminus of the double strand may be modified without significant loss of activity.

The double stranded RNA molecule of the invention may be prepared by a method that includes the following steps:

(a) synthesizing two RNA strands each having a length from 19-25, *e.g.*, from 19-23  
25 nucleotides, wherein said RNA strands are capable of forming a double stranded RNA molecule, wherein preferably at least one strand has a 3'-overhang from 1-5 nucleotides; and

(b) combining the synthesized RNA strands under conditions, where a double stranded RNA molecule is formed, which is capable of mediating target-specific nucleic acid modifications, particularly RNAi interference and/or DNA methylation.

Methods of synthesizing RNA molecules are known in the art. In this context, it is particularly referred to chemical synthesis methods as described in Verma and Eckstein (1998).

5 The single-stranded RNAs can also be prepared by enzymatic transcription from synthetic DNA templates or from DNA plasmids isolated from recombinant bacteria. Typically, phage RNA polymerases are used such as T7, T3 or SP6 RNA polymerase (Milligan and Uhlenbeck, 1989).

Another aspect of the present invention relates to a method of mediating target-specific nucleic acid modifications, particularly RNA interference and/or DNA methylation  
10 in a cell or an organism. The method in can include the following steps:

- (a) contacting the cell or organism with the double stranded RNA molecule of the invention under conditions wherein target-specific nucleic acid modifications can occur; and
- (b) mediating a target-specific nucleic acid modification effected by the double stranded RNA towards a target nucleic acid having a sequence portion substantially  
15 corresponding to the double stranded RNA.

The contacting step preferably includes (a) introducing the double stranded RNA molecule into a target cell, *e.g.*, an isolated target cell, *e.g.*, in cell culture, a unicellular microorganism or a target cell or a plurality of target cells within a multicellular organism. More preferably, the introducing step comprises a carrier-mediated delivery, *e.g.*, by  
20 liposomal carriers or by injection.

The method of the invention may be used for determining the function of a gene in a cell or an organism (or for modulating the function of a gene in a cell or an organism) capable of mediating RNA interference. The cell can be a eukaryotic cell or a cell line, *e.g.*, a plant cell or an animal cell, such as a mammalian cell, *e.g.*, an embryonic cell, a pluripotent  
25 stem cell, a tumor cell, *e.g.*, a teratocarcinoma cell or a virus-infected cell. The organism is preferably a eukaryotic organism, *e.g.*, a plant or an animal, such as a mammal, particularly a human.

The target gene to which the RNA molecule of the invention is directed can be associated with a pathological condition. For example, the gene may be a pathogen-associated gene, *e.g.*, a viral gene, a tumor-associated gene or an autoimmune disease-associated gene. The target gene may also be a heterologous gene expressed in a  
30

recombinant cell or a genetically altered organism. By modulating, particularly, inhibiting the function of such a gene, valuable information and therapeutic benefits in the agricultural field, or in the medicine or veterinary medicine field, may be obtained.

The dsRNA is usually administered as a pharmaceutical composition. The  
5 administration may be carried out by known methods, wherein a nucleic acid is introduced into a desired target cell *in vitro* or *in vivo*. Commonly used gene transfer techniques include calcium phosphate, DEAE-dextran, electroporation, microinjection and viral methods (Graham, F. L. and van der Eb, A. J. (1973) *Virology* 52, 456; McCutchan, J. H. and Pagano, J. S. (1968), *J. Natl. Cancer Inst.* 41, 351; Chu, G. *et al* (1987), *Nucl. Acids Res.* 15, 1311;  
10 Fraley, R. *et al.* (1980), *J. Biol. Chem.* 255, 10431; Capecchi, M.R. (1980), *Cell* 22, 479). A recent addition to this arsenal of techniques for the introduction of DNA into cells is the use of cationic liposomes (Feigner, P. L. *et al.* (1987), *Proc. Natl. Acad. Sci USA* 84, 7413). Commercially available cationic lipid formulations are, *e.g.*, Tfx 50 (Promega) or Lipofectamin2000 (Life Technologies). Thus, the invention also relates to a pharmaceutical  
15 composition containing as an active agent at least one double stranded RNA molecule as described above and a pharmaceutical carrier. The composition may be used for diagnostic and for therapeutic applications in human medicine or in veterinary medicine.

For diagnostic or therapeutic applications, the composition may be in the form of a solution, *e.g.*, an injectable solution, a cream, ointment, tablet, suspension or the like. The  
20 composition may be administered in any suitable way, *e.g.*, by injection, by oral, topical, nasal, rectal application, *etc.* The carrier may be any suitable pharmaceutical carrier.

Preferably, a carrier is used, which is capable of increasing the efficacy of the RNA molecules to enter the target-cells. Suitable examples of such carriers are liposomes, particularly cationic liposomes. A further preferred administration method is injection.

25 The RNAi method described herein can also be used for the functional analysis of eukaryotic cells, or eukaryotic non-human organisms, preferably mammalian cells or organisms and most preferably human cells, *e.g.*, cell lines such as HeLa or 293 or rodents, *e.g.*, rats and mice. By transfection with suitable double stranded RNA molecules which are homologous to a predetermined target gene or DNA molecules encoding a suitable double  
30 stranded RNA molecule a specific knockout phenotype can be obtained in a target cell, *e.g.*, in cell culture or in a target organism. Surprisingly it was found that the presence of short

double stranded RNA molecules do not result in an interferon response from the host cell or host organism.

Thus, a further subject matter of the invention is a eukaryotic cell or a eukaryotic non-human organism exhibiting a target gene-specific knockout phenotype comprising an at least partially deficient expression of at least one endogenous target gene wherein said cell or organism is transfected with at least one double stranded RNA molecule capable of inhibiting the expression of at least one endogenous target gene or with a DNA encoding at least one double stranded RNA molecule capable of inhibiting the expression of at least one endogenous target gene. It should be noted that the present invention allows a target-specific knockout of several different endogenous genes due to the specificity of RNAi.

Gene-specific knockout phenotypes of cells or non-human organisms, particularly of human cells or non-human mammals may be used in analytic procedures, *e.g.*, in the functional and/or phenotypical analysis of complex physiological processes such as analysis of gene expression profiles and/or proteomes. For example, one may prepare the knock-out phenotypes of human genes in cultured cells which are assumed to be regulators of alternative splicing processes. Among these genes are particularly the members of the SR splicing factor family, *e.g.*, ASF/SF2, SC35, SRp20, SRp40 or SRp55. Further, the effect of SR proteins on the mRNA profiles of predetermined alternatively spliced genes such as CD44 may be analyzed. Preferably the analysis is carried out by high-throughput methods using oligonucleotide based chips.

Using RNAi-based knockout technologies, the expression of an endogenous target gene may be inhibited in a target cell or a target organism. The endogenous gene may be complemented by an exogenous target nucleic acid coding for the target protein or a variant or mutated form of the target protein, *e.g.*, a gene or cDNA, a which may optionally be fused to a further nucleic acid sequence encoding a detectable peptide or polypeptide, *e.g.*, an affinity tag, particularly a multiple affinity tag. Variants or mutated forms of the target gene differ from the endogenous target gene in that they encode a gene product which differs from the endogenous gene product on the amino acid level by substitutions, insertions and/or deletions of single or multiple amino acids. The variants or mutated forms may have the same biological activity as the endogenous target gene. On the other hand, the variant or mutated target gene may also have a biological activity, which differs from the biological

activity of the endogenous target gene, *e.g.*, a partially deleted activity, a completely deleted activity, an enhanced activity, *etc.*

The complementation may be accomplished by coexpressing the polypeptide encoded by the exogenous nucleic acid, *e.g.*, a fusion protein comprising the target protein and the affinity tag and the double stranded RNA molecule for knocking out the endogenous gene in the target cell. This coexpression may be accomplished by using a suitable expression vector expressing both the polypeptide encoded by the exogenous nucleic acid, *e.g.*, the tag-modified target protein and the double stranded RNA molecule or alternatively by using a combination of expression vectors. Proteins and protein complexes which are synthesized de novo in the target cell will contain the exogenous gene product, *e.g.*, the modified fusion protein. In order to avoid suppression of the exogenous gene product expression by the RNAi duplex molecule, the nucleotide sequence encoding the exogenous nucleic acid may be altered, on the DNA level (with or without causing mutations on the amino acid level) in the part of the sequence which is homologous to the double stranded RNA molecule. Alternatively, the endogenous target gene may be complemented by corresponding nucleotide sequences from other species, *e.g.*, from mouse.

Preferred applications for the cell or organism of the invention is the analysis of gene expression profiles and/or proteomes. In an especially preferred embodiment, an analysis of a variant or mutant form of one or several target proteins is carried out, wherein said variant or mutant forms are reintroduced into the cell or organism by an exogenous target nucleic acid as described above. The combination of knockout of an endogenous gene and rescue by using mutated, *e.g.*, partially deleted exogenous target has advantages compared to the use of a knockout cell. Further, this method is particularly suitable for identifying functional domains of the target protein. In a further preferred embodiment a comparison, *e.g.*, of gene expression profiles and/or proteomes and/or phenotypic characteristics of at least two cells or organisms is carried out. These organisms are selected from:

- (i) a control cell or control organism without target gene inhibition,
- (ii) a cell or organism with target gene inhibition, and
- (iii) a cell or organism with target gene inhibition plus target gene complementation by an exogenous target nucleic acid.

The method and cells of the invention are also suitable in a procedure for identifying and/or characterizing pharmacological agents, *e.g.*, identifying new pharmacological agents from a collection of test substances and/or characterizing mechanisms of action and/or side effects of known pharmacological agents.

Thus, the present invention also relates to a system for identifying and/or characterizing pharmacological agents acting on at least one target protein comprising:

(a) a eukaryotic cell or a eukaryotic non-human organism capable of expressing at least one endogenous target gene coding for said target protein,

(b) at least one double stranded RNA molecule capable of inhibiting the expression of said at least one endogenous target gene, and

(c) a test substance or a collection of test substances wherein pharmacological properties of said test substance or said collection are to be identified and/or characterized.

Further, the system as described above preferably comprises:

(d) at least one exogenous target nucleic acid coding for the target protein or a variant or mutated form of the target protein wherein said exogenous target nucleic acid differs from the endogenous target gene on the nucleic acid level such that the expression of the exogenous target nucleic acid is substantially less inhibited by the double stranded RNA molecule than the expression of the endogenous target gene.

Furthermore, the RNA knockout complementation method may be used for preparative purposes, *e.g.*, for the affinity purification of proteins or protein complexes from eukaryotic cells, particularly mammalian cells and more particularly human cells. In this embodiment of the invention, the exogenous target nucleic acid preferably codes for a target protein which is fused to an affinity tag.

The preparative method may be employed for the purification of high molecular weight protein complexes which preferably have a mass of 150 kD and more preferably of 500 kD and which optionally may contain nucleic acids such as RNA. Specific examples are the heterotrimeric protein complex consisting of the 20 kD, 60 kD and 90 kD proteins of the U4/U6 snRNP particle, the splicing factor SF3b from the 17S U2 snRNP consisting of 5 proteins having molecular weights of 14, 49, 120, 145 and 155 kD and the 25S U4/U6/U5 tri-snRNP particle containing the U4, U5 and U6 snRNA molecules and about 30 proteins, which has a molecular weight of about 1.7 MD.

This method is suitable for functional proteome analysis in mammalian cells, particularly human cells.

The invention is further illustrated by the following examples, which should not be construed as further limiting. The contents of all references, pending patent applications and published patents, cited throughout this application are hereby expressly incorporated by reference.

## EXAMPLES

### Example 1. RNA Interference Mediated by Small Synthetic RNAs

#### 1.1. Experimental Procedures

##### 1.1.1. *In vitro* RNAi

*In vitro* RNAi and lysate preparations were performed as described previously (Tuschl *et al.*, 1999; Zamore *et al.*, 2000). It is critical to use freshly dissolved creatine kinase (Roche) for optimal ATP regeneration. The RNAi translation assays (Fig. 1) were performed with dsRNA concentrations of 5 nM and an extended pre-incubation period of 15 min at 25°C prior to the addition of *in vitro* transcribed, capped and polyadenylated Pp-luc and Rr-luc reporter mRNAs. The incubation was continued for 1 h and the relative amount of Pp-luc and Rr-luc protein was analyzed using the dual luciferase assay (Promega) and a Monolight 3010C luminometer (PharMingen).

##### 1.1.2. RNA Synthesis.

Standard procedures were used for *in vitro* transcription of RNA from PCR templates carrying T7 or SP6 promoter sequences, see for example (Tuschl *et al.*, 1998). Synthetic RNA was prepared using Expedite RNA phosphoramidites (Proligo). The 3' adapter oligonucleotide was synthesized using dimethoxytrityl-1,4-benzenedimethanol-succinyl aminopropyl-CPG. The oligoribonucleotides were deprotected in 3 ml of 32% ammonia/ethanol (3/1) for 4 h 55°C at (Expedite RNA) or 16 h at 55°C (3' and 5' adapter DNA/RNA chimeric oligonucleotides) and then desialylated and gel-purified as described previously (Tuschl *et al.*, 1993). RNA transcripts for dsRNA preparation including long 3' overhangs were generated from PCR templates that contained a T7 promoter in sense and an SP6 promoter in antisense direction. The transcription template for sense and antisense target

RNA was PCR-amplified with  
GCGTAATACGACTCACTATAGAAACAATTGCTTTTACAG (underlined, T7 promoter) as  
5' primer and ATTTAGGTGACACTATAGGCATAAAGAATTGAAGA (underlined, SP6  
promoter) as 3' primer and the linearized Pp-luc plasmid (PGEM-luc sequence) (Tuschl *et*  
5 *al.*, 1999) as template; the T7-transcribed sense RNA was 177 nt long with the Pp-luc  
sequence between pos. 113-273 relative to the start codon and followed by 17 nt of the  
complement of the SP6 promoter sequence at the 3' end. Transcripts for blunt-ended dsRNA  
formation were prepared by transcription from two different PCR products which only  
contained a single promoter sequence.

10 DsRNA annealing was carried out using a phenol/chloroform extraction. Equimolar  
concentration of sense and antisense RNA (50 nM to 10  $\mu$ M, depending on the length and  
amount available) in 0.3 M NaOAc (pH 6) were incubated for 30 s at 90°C and then  
extracted at room temperature with an equal volume of phenol/chloroform, and followed by a  
chloroform extraction to remove residual phenol. The resulting dsRNA was precipitated by  
15 addition of 2.5-3 volumes of ethanol. The pellet was dissolved in lysis buffer (100 mM KCl,  
30 mM HEPES-KOH, pH 7.4, 2 mM Mg(OAc)<sub>2</sub>) and the quality of the dsRNA was verified  
by standard agarose gel electrophoreses in 1 x TAE-buffer. The 52 bp dsRNAs with the 17  
nt and 20 nt 3' overhangs (FIG. 6) were annealed by incubating for 1 min at 95°C, then  
rapidly cooled to 70°C and followed by slow cooling to room temperature over a 3 h period  
20 (50  $\mu$ l annealing reaction, 1  $\mu$ M strand concentration, 300 mM NaCl, 10 mM Tris-HCl, pH  
7.5). The dsRNAs were then phenol/chloroform extracted, ethanol-precipitated and  
dissolved in lysis buffer.

Transcription of internally <sup>32</sup>P -radiolabeled RNA used for dsRNA preparation (FIGs.  
2 and 4) was performed using 1 mM ATP, CTP, GTP, 0.1 or 0.2 mM UTP, and 0.2-0.3  $\mu$ M  
25 32P-UTP (3000 Ci/mmol), or the respective ratio for radiolabeled nucleoside triphosphates  
other than UTP. Labeling of the cap of the target RNAs was performed as described  
previously. The target RNAs were gel-purified after cap-labeling.

### 1.1.3. Cleavage Site Mapping

30 Standard RNAi reactions were performed by pre-incubating 10 nM dsRNA for 15  
min followed by addition of 10 nM cap-labeled target RNA. The reaction was stopped after



a further 2 h (FIG. 2A) or 2.5 h incubation (FIG. 5B and 6B) by proteinase K treatment (Tuschl *et al.*, 1999). The samples were then analyzed on 8 or 10% sequencing gels. The 21 and 22 nt synthetic RNA duplexes were used at 100 nM final concentration (Fig. 5B).

#### 1.1.4. Cloning of ~21 nt RNAs

The 21 nt RNAs were produced by incubation of radiolabeled dsRNA in *Drosophila* lysate in absence of target RNA (200  $\mu$ l reaction, 1 h incubation, 50 nM dsP111, or 100 nM dsP52 or dsP39). The reaction mixture was subsequently treated with proteinase K (Tuschl *et al.*, 1999) and the dsRNA-processing products were separated on a denaturing 15% polyacrylamide gel. A band, including a size range of at least 18 to 24 nt, was excised, eluted into 0.3 M NaCl overnight at 4°C and in siliconized tubes. The RNA was recovered by ethanol-precipitation and dephosphorylated (30  $\mu$ l reaction, 30 min, 50°C, 10 U alkaline phosphatase, Roche). The reaction was stopped by phenol/chloroform extraction and the RNA was ethanol-precipitated. The 3' adapter oligonucleotide (pUUUaaccgcatccttctcx: uppercase, RNA; lowercase, DNA; p, phosphate; x, 4-hydroxymethylbenzyl) was then ligated to the dephosphorylated ~21 nt RNA (20  $\mu$ l reaction, 30 min, 37°C, 5  $\mu$ M 3' adapter, 50 mM Tris-HCl, pH 7.6, 10 mM MgCl<sub>2</sub>, 0.2 mM ATP, 0.1 mg/ml acetylated BSA, 15% DMSO, 25 U T4 RNA ligase, Amersham-Pharmacia)(Pan and Uhlenbeck, 1992). The ligation reaction was stopped by the addition of an equal volume of 8 M urea/50 mM EDTA stop mix and directly loaded on a 15% gel. Ligation yields were greater than 50%. The ligation product was recovered from the gel and 5'-phosphorylated(20  $\mu$ l reaction, 30 min, 37°C, 2 mM ATP, 5 U T4 polynucleotide kinase, NEB). The phosphorylation reaction was stopped by phenol/chloroform extraction and RNA was recovered by ethanol-precipitation. Next, the 5' adapter (tactaatacgactactAAA: uppercase, RNA; lowercase, DNA) was ligated to the phosphorylated ligation product as described above. The new ligation product was gel-purified and eluted from the gel slice in the presence of reverse transcription primer (GACTAGCTGGAATTCAAGGATGCGGTAAA: bold, Eco RI site) used as carrier. Reverse transcription (15  $\mu$ l reaction, 30 min, 42°C, 150 U Superscript II reverse transcriptase, Life Technologies) was followed by PCR using as 5' primer CAGCCAACGGAATTCATACGACTCACTAAA (bold, Eco RI site) and the 3' RT primer. The PCR product was purified by phenol/chloroform extraction and ethanol precipitated.

The PCR product was then digested with Eco RI (NEB) and concatamerized using T4 DNA ligase (high cont., NEB). Concatamers of a size range of 200 to 800 bp were separated on a low-melt agarose gel, recovered from the gel by a standard melting and phenol extraction procedure, and ethanol-precipitated. The unpaired ends were filled in by incubation with Taq polymerase under standard conditions for 15 min at 72°C and the DNA product was directly ligated into the pCR2.1-TOPO vector using the TOPO TA cloning kit (Invitrogen). Colonies were screened using PCR and M13-20 and M13 Reverse sequencing primers. PCR products were directly submitted for custom sequencing (Sequence Laboratories Göttingen GmbH, Germany). On average, four to five 21mer sequences were obtained per clone.

#### 1.1.5. 2D-TLC Analysis

Nuclease P1 digestion of radiolabeled, gel-purified siRNAs and 2D-TLC was carried out as described (Zamore *et al.*, 2000). Nuclease T2 digestion was performed in 10  $\mu$ l reactions for 3 h at 50°C in 10 mM ammonium acetate (pH 4.5) using 2  $\mu$ g/ $\mu$ l carrier tRNA and 30 U ribonuclease T2 (Life Technologies). The migration of non-radioactive standards was determined by UV shadowing. The identity of nucleoside-3',5'-diphosphates was confirmed by co-migration of the T2 digestion products with standards prepared by 5'-<sup>32</sup>P-phosphorylation of commercial nucleoside 3'-monophosphates using  $\gamma$ -<sup>32</sup>P-ATP and T4 polynucleotide kinase (data not shown).

### 1.2. Results and Discussion

#### 1.2.1. Length Requirements for Processing of dsRNA to 21 and 22 nt RNA

##### Fragments

Lysate prepared from *D. melanogaster* syncytial embryos recapitulates RNAi *in vitro* providing a novel tool for biochemical analysis of the mechanism of RNAi (Tuschl *et al.*, 1999; Zamore *et al.*, 2000). *In vitro* and *in vivo* analysis of the length requirements of dsRNA for RNAi has revealed that short dsRNA (<150 bp) are less effective than longer dsRNAs in degrading target mRNA (Caplen *et al.*, 2000; Hammond *et al.*, 2000; Ngo *et al.*, 1998; Tuschl *et al.*, 1999). The reasons for reduction in mRNA degrading efficiency are not understood. We therefore examined the precise length requirement of dsRNA for target RNA degradation under optimized conditions in the *Drosophila* lysate (Zamore *et al.*, 2000).

Several series of dsRNAs were synthesized and directed against firefly luciferase (Pp-luc) reporter RNA. The specific suppression of target RNA expression was monitored by the dual luciferase assay (Tuschl *et al.*, 1999) (FIGs. 1A and 1B). We detected specific inhibition of target RNA expression for dsRNAs as short as 38 bp, but dsRNAs of 29 to 36 bp were not effective in this process. The effect was independent of the target position and the degree of inhibition of Pp-luc mRNA expression correlated with the length of the dsRNA, *i.e.*, long dsRNAs were more effective than short dsRNAs.

It has been suggested that the 21-23 nt RNA fragments generated by processing of dsRNAs are the mediators of RNA interference and co-suppression (Hamilton and Baulcombe, 1999; Hammond *et al.*, 2000; Zamore *et al.*, 2000). We therefore analyzed the rate of 21-23 nt fragment formation for a subset of dsRNAs ranging in size between 501 to 29 bp. Formation of 21-23 nt fragments in *Drosophila* lysate (FIG. 2) was readily detectable for 39 to 501 bp long dsRNAs but was significantly delayed for the 29 bp dsRNA. This observation is consistent with a role of 21-23 nt fragments in guiding mRNA cleavage and provides an explanation for the lack of RNAi by 30 bp dsRNAs. The length dependence of 21-23 mer formation is likely to reflect a biologically relevant control mechanism to prevent the undesired activation of RNAi by short intramolecular base-paired structures of regular cellular RNAs.

#### 1.2.2. 39 bp dsRNA Mediates Target RNA Cleavage at a Single Site

Addition of dsRNA and 5'-capped target RNA to the *Drosophila* lysate results in sequence-specific degradation of the target RNA (Tuschl *et al.*, 1999). The target mRNA is only cleaved within the region of identity with the dsRNA and many of the target cleavage sites were separated by 21-23 nt (Zamore *et al.*, 2000). Thus, the number of cleavage sites for a given dsRNA was expected to roughly correspond to the length of the dsRNA divided by 21. We mapped the target cleavage sites on a sense and an antisense target RNA which was 5' radiolabeled at the cap (Zamore *et al.*, 2000) (FIG.s 3A and 3B). Stable 5' cleavage products were separated on a sequencing gel and the position of cleavage was determined by comparison with a partial RNase T1 and an alkaline hydrolysis ladder from the target RNA.

Consistent with the previous observation (Zamore *et al.*, 2000), all target RNA cleavage sites were located within the region of identity to the dsRNA. The sense or the

antisense target was only cleaved once by 39 bp dsRNA. Each cleavage site was located 10 nt from the 5' end of the region covered by the dsRNA (FIG. 3B). The 52 bp dsRNA, which shares the same 5' end with the 39 bp dsRNA, produces the same cleavage site on the sense target, located 10 nt from the 5' end of the region of identity with the dsRNA, in addition to two weaker cleavage sites 23 and 24 nt downstream of the first site. The antisense target was only cleaved once, again 10 nt from the 5' end of the region covered by its respective dsRNA. Mapping of the cleavage sites for the 38 to 49 bp dsRNAs shown in FIG. 1 showed that the first and predominant cleavage site was always located 7 to 10 nt downstream of the region covered by the dsRNA (data not shown). This suggests that the point of target RNA cleavage is determined by the end of the dsRNA and could imply that processing to 21-23 mers starts from the ends of the duplex.

Cleavage sites on sense and antisense target for the longer 111 bp dsRNA were much more frequent than anticipated and most of them appear in clusters separated by 20 to 23 nt (FIG.s 3A and 3B). As for the shorter dsRNAs, the first cleavage site on the sense target is 10 nt from the 5' end of the region spanned by the dsRNA, and the first cleavage site on the antisense target is located 9 nt from the 5' end of the region covered by the dsRNA. It is unclear what causes this disordered cleavage, but one possibility could be that longer dsRNAs may not only get processed from the ends but also internally, or there are some specificity determinants for dsRNA processing which we do not yet understand. Some irregularities to the 21-23 nt spacing were also previously noted (Zamore *et al.*, 2000). To better understand the molecular basis of dsRNA processing and target RNA recognition, we decided to analyze the sequences of the 21-23 nt fragments generated by processing of 39, 52, and 111 bp dsRNAs in the *Drosophila* lysate.

### 1.2.3 dsRNA is Processed to 21 and 22 nt RNAs by an RNase III-Like Mechanism

In order to characterize the 21-23 nt RNA fragments, we examined the 5' and 3' termini of the RNA fragments. Periodate oxidation of gel-purified 21-23 nt RNAs followed by  $\beta$ -elimination indicated the presence of a terminal 2' and 3' hydroxyl groups. The 21-23 mers were also responsive to alkaline phosphatase treatment indicating the presence of a 5' terminal phosphate group. The presence of 5' phosphate and 3' hydroxyl termini suggests

that the dsRNA could be processed by an enzymatic activity similar to *E. coli* RNase III (for reviews, see Dunn, 1982; Nicholson, 1999; Robertson, 1990; Robertson, 1982).

Directional cloning of 21-23 nt RNA fragments was performed by ligation of a 3' and 5' adapter oligonucleotide to the purified 21-23 mers using T4 RNA ligase. The ligation products were reverse transcribed, PCR-amplified, concatamerized, cloned, and sequenced. Over 220 short RNAs were sequenced from dsRNA processing reactions of the 39, 52 and 111 bp dsRNAs (FIG. 4A). We found the following length distribution: 1% 18 nt, 5% 19 nt, 12% 20 nt, 45% 21 nt, 28% 22 nt, 6% 23 nt, and 2% 24 nt. Sequence analysis of the 5' terminal nucleotide of the processed fragments indicated that oligonucleotides with a 5' guanosine were underrepresented. This bias was most likely introduced by T4 RNA ligase which discriminates against 5' phosphorylated guanosine as donor oligonucleotide; no significant sequence bias was seen at the 3' end. Many of the ~21 nt fragments derived from the 3' ends of the sense or antisense strand of the duplexes include 3' nucleotides that are derived from untemplated addition of nucleotides during RNA synthesis using T7 RNA polymerase. Interestingly, a significant number of endogenous *Drosophila* ~21 nt RNAs were also cloned, some of them from LTR and non-LTR retrotransposons (data not shown). This is consistent with a possible role for RNAi in transposon silencing.

The ~21 nt RNAs appear in clustered groups (FIG. 4A) which cover the entire dsRNA sequences. Apparently, the processing reaction cuts the dsRNA by leaving staggered 3' ends, another characteristic of RNase III cleavage. For the 39 bp dsRNA, two clusters of ~21 nt RNAs were found from each dsRNA-constituting strand including overhanging 3' ends, yet only one cleavage site was detected on the sense and antisense target (FIG.s 3A and 3B). If the ~21 nt fragments were present as single-stranded guide RNAs in a complex that mediates mRNA degradation, it could be assumed that at least two target cleavage sites exist, but this was not the case. This suggests that the ~21 nt RNAs may be present in double stranded form in the endonuclease complex but that only one of the strands can be used for target RNA recognition and cleavage. The use of only one of the ~21 nt strands for target cleavage may simply be determined by the orientation in which the ~21 nt duplex is bound to the nuclease complex. This orientation is defined by the direction in which the original dsRNA was processed.

The ~21 mer clusters for the 52 bp and 111 bp dsRNA are less well defined when compared to the 39 bp dsRNA. The clusters are spread over regions of 25 to 30 nt most likely representing several distinct subpopulations of ~21 nt duplexes and therefore guiding target cleavage at several nearby sites. These cleavage regions are still predominantly separated by 20 to 23 nt intervals. The rules determining how regular dsRNA can be processed to ~21 nt fragments are not yet understood, but it was previously observed that the approx. 21-23 nt spacing of cleavage sites could be altered by a run of uridines (Zamore *et al.*, 2000). The specificity of dsRNA cleavage by *E. coli* RNase III appears to be mainly controlled by antideterminants, *i.e.*, excluding some specific base-pairs at given positions relative to the cleavage site (Zhang and Nicholson, 1997).

To test whether sugar-, base- or cap-modification were present in processed ~21 nt RNA fragments, we incubated radiolabeled 505 bp Pp-luc dsRNA in lysate for 1 h, isolated the ~21 nt products, and digested it with P1 or T2 nuclease to mononucleotides. The nucleotide mixture was then analyzed by 2D thin-layer chromatography (FIG. 4B). None of the four natural ribonucleotides were modified as indicated by P1 or T2 digestion. We have previously analyzed adenosine to inosine conversion in the ~21 nt fragments (after a 2 h incubation) and detected a small extent (<0.7%) deamination (Zamore *et al.*, 2000); shorter incubation in lysate (1 h) reduced this inosine fraction to barely detectable levels. RNase T2, which cleaves 3' of the phosphodiester linkage, produced nucleoside 3'-phosphate and nucleoside 3', 5'-diphosphate, thereby indicating the presence of a 5'-terminal monophosphate. All four nucleoside 3', 5'-diphosphates were detected and suggest that the internucleotidic linkage was cleaved with little or no sequence-specificity. In summary, the ~21 nt fragments are unmodified and were generated from dsRNA such that 5'-monophosphates and 3'-hydroxyls were present at the 5'-end.

#### 1.2.4 Synthetic 21 and 22 nt RNAs Mediate Target RNA Cleavage

Analysis of the products of dsRNA processing indicated that the ~21 nt fragments are generated by a reaction with all the characteristics of an RNase III cleavage reaction (Dunn, 1982; Nicholson, 1999; Robertson, 1990; Robertson, 1982). RNase III makes two staggered cuts in both strands of the dsRNA, leaving a 3' overhang of about 2 nt. We chemically synthesized 21 and 22 nt RNAs, identical in sequence to some of the cloned ~21 nt

fragments, and tested them for their ability to mediate target RNA degradation (FIG.s 5A and 5B). The 21 and 22 nt RNA duplexes were incubated at 100 nM concentrations in the lysate, a 10-fold higher concentration than the 52 bp control dsRNA. Under these conditions, target RNA cleavage is readily detectable. Reducing the concentration of 21 and 22 nt duplexes  
5 from 100 to 10 nM does still cause target RNA cleavage. Increasing the duplex concentration from 100 nM to 1000 nM however does not further increase target cleavage, probably due to a limiting protein factor within the lysate.

In contrast to 29 or 30 bp dsRNAs that did not mediate RNAi, the 21 and 22 nt dsRNAs with overhanging 3' ends of 2 to 4 nt mediated efficient degradation of target RNA  
10 (duplexes 1, 3, 4, 6, FIG.s 5A and 5B). Blunt-ended 21 or 22 nt dsRNAs (duplexes 2, 5, and 7, FIG.s 5A and 5B) were reduced in their ability to degrade the target and indicate that overhanging 3' ends are critical for reconstitution of the RNA-protein nuclease complex. The single-stranded overhangs may be required for high affinity binding of the ~21 nt duplex to the protein components. A 5' terminal phosphate, although present after dsRNA processing,  
15 was not required to mediate target RNA cleavage and was absent from the short synthetic RNAs.

The synthetic 21 and 22 nt duplexes guided cleavage of sense as well as antisense targets within the region covered by the short duplex. This is an important result considering that a 39 bp dsRNA, which forms two pairs of clusters of ~21 nt fragments (Fig. 2), cleaved  
20 sense or antisense targets only once and not twice. We interpret this result by suggesting that only one of two strands present in the ~21 nt duplex is able to guide target RNA cleavage and that the orientation of the ~21 nt duplex in the nuclease complex is determined by the initial direction of dsRNA processing. The presentation of an already perfectly processed ~21 nt duplex to the *in vitro* system however does allow formation of the active sequence specific  
25 nuclease complex with two possible orientations of the symmetric RNA duplex. This results in cleavage of sense as well as antisense target within the region of identity with the 21 nt RNA duplex.

The target cleavage site is located 11 or 12 nt downstream of the first nucleotide that is complementary to the 21 or 22 nt guide sequence, *i.e.*, the cleavage site is near the center  
30 of the region covered by the 21 or 22 nt RNAs (FIG.s 4A and 4B). Displacing the sense strand of a 22 nt duplex by two nucleotides (compare duplexes 1 and 3 in FIG. 5A) displaced

the cleavage site of only the antisense target by two nucleotides. Displacing both sense and antisense strands by two nucleotides shifted both cleavage sites by two nucleotides (compare duplexes 1 and 4). It should be possible to design a pair of 21 or 22 nt RNAs to cleave a target RNA at almost any given position.

5       The specificity of target RNA cleavage guided by 21 and 22 nt RNAs appears exquisite as no aberrant cleavage sites are detected (Fig. 5B). It should however be noted, that the nucleotides present in the 3' overhang of the 21 and 22 nt RNA duplex may contribute less to substrate recognition than the nucleotides near the cleavage site. This is based on the observation that the 3' most nucleotide in the 3' overhang of the active duplexes 10       1 or 3 (FIG. 5A) is not complementary to the target. A detailed analysis of the specificity of RNAi can now be readily undertaken using synthetic 21 and 22 nt RNAs.

Based on the evidence that synthetic 21 and 22 nt RNAs with overhanging 3' ends mediate RNA interference, we propose to name the ~21 nt RNAs "short interfering RNAs" or siRNAs and the respective RNA-protein complex a "small interfering ribonucleoprotein 15       particle," or siRNP.

#### 1.2.5. 3' Overhangs of 20 nt on short dsRNAs inhibit RNAi

We have shown that short blunt-ended dsRNAs appear to be processed from the ends of the dsRNA. During our study of the length dependence of dsRNA in RNAi, we have also 20       analyzed dsRNA with 17 to 20 nt overhanging 3' ends and found to our surprise that they were less potent than blunt-ended dsRNAs. The inhibitory effect of long 3' ends was particularly pronounced for dsRNAs up to 100 bp but was less dramatic for longer dsRNAs. The effect was not due to imperfect dsRNA formation based on native gel analysis (data not shown). We tested if the inhibitory effect of long overhanging 3' ends could be used as a tool 25       to direct dsRNA processing to only one of the two ends of a short RNA duplex.

We synthesized four combinations of the 52 bp model dsRNA, blunt-ended, 3' extension on only the sense strand, 3'-extension on only the antisense strand, and double 3' extension on both strands, and mapped the target RNA cleavage sites after incubation in lysate (FIG.s 6A and 6B). The first and predominant cleavage site of the sense target was 30       lost when the 3' end of the antisense strand of the duplex was extended, and vice versa, the strong cleavage site of the antisense target was lost when the 3' end of sense strand of the



duplex was extended. 3' Extensions on both strands rendered the 52 bp dsRNA virtually inactive. One explanation for the dsRNA inactivation by ~20 nt 3' extensions could be the association of single-stranded RNA-binding proteins which could interfere with the association of one of the dsRNA-processing factors at this end. This result is also consistent with our model where only one of the strands of the siRNA duplex in the assembled siRNP is able to guide target RNA cleavage. The orientation of the strand that guides RNA cleavage is defined by the direction of the dsRNA processing reaction. It is likely that the presence of 3' staggered ends may facilitate the assembly of the processing complex. A block at the 3' end of the sense strand will only permit dsRNA processing from the opposing 3' end of the antisense strand. This in turn generates siRNP complexes in which only the antisense strand of the siRNA duplex is able to guide sense target RNA cleavage. The same is true for the reciprocal situation.

The less pronounced inhibitory effect of long 3' extensions in the case of longer dsRNAs (500 bp, data not shown) suggests to us that long dsRNAs may also contain internal dsRNA-processing signals or may get processed cooperatively due to the association of multiple cleavage factors.

#### 1.2.6. A Model for dsRNA-Directed mRNA Cleavage

The new biochemical data update the model for how dsRNA targets mRNA for destruction (FIG. 7). Double stranded RNA is first processed to short RNA duplexes of predominantly 21 and 22 nt in length and with staggered 3' ends similar to an RNase III-like reaction (Dunn, 1982.; Nicholson, 1999; Robertson, 1982). Based on the 21-23 nt length of the processed RNA fragments it has already been speculated that an RNase III-like activity may be involved in RNAi (Bass, 2000). This hypothesis is further supported by the presence of 5' phosphates and 3' hydroxyls termini at the termini of the siRNAs as observed in RNase III reaction products (Dunn, 1982; Nicholson, 1999). Bacterial RNase III and the eukaryotic homologs Rnt1p in *S. cerevisiae* and Pac1p in *S. pombe* have been shown to function in processing of ribosomal RNA as well as snRNA and snoRNAs (see for example Chanfreau *et al.*, 2000).

Little is known about the biochemistry of RNase III homologs from plants, or animals, including humans. Two families of RNase III enzymes have been identified

predominantly by database-guided sequence analysis or cloning of cDNAs. The first RNase III family is represented by the 1327 amino acid long *D. melanogaster* protein drosha (Acc. AF116572). The C-terminus is composed of two RNase III and one dsRNA-binding domain and the N-terminus is of unknown function. Close homologs are also found in *C. elegans* (Acc. AF160248) and human (Acc. AF189011) (Filippov *et al.*, 2000; Wu *et al.*, 2000). The drosha-like human RNase III was recently cloned and characterized (Wu *et al.*, 2000). The gene is ubiquitously expressed in human tissues and cell lines, and the protein is localized in the nucleus and the nucleolus of the cell. Based on results inferred from antisense inhibition studies, a role of this protein for rRNA processing was suggested. The second class is represented by the *C. elegans* gene K12H4.8 (Acc. S44849) coding for a 1822 amino acid long protein. This protein has an N-terminal RNA helicase motif which is followed by 2 RNase III catalytic domains and a dsRNA-binding motif, similar to the drosha RNase III family. There are close homologs in *S. pombe* (Acc. Q09884), *A. thaliana* (Acc. AF187317), *D. melanogaster* (Acc. AE003740), and human (Acc. AB028449) (Filippov *et al.*, 2000; Jacobsen *et al.*, 1999; Matsuda *et al.*, 2000). Possibly the K12H4.8 RNase III/helicase is the likely candidate to be involved in RNAi.

Genetic screens in *C. elegans* identified *rde-1* and *rde-4* as essential for activation of RNAi without an effect on transposon mobilization or co-suppression (Dernburg *et al.*, 2000; Grishok *et al.*, 2000; Ketting and Plasterk, 2000; Tabara *et al.*, 1999). This led to the hypothesis that these genes are important for dsRNA processing but are not involved in mRNA target degradation. The function of both genes is as yet unknown, the *rde-1* gene product is a member of a family of proteins similar to the rabbit protein eIF2C (Tabara *et al.*, 1999), and the sequence of *rde-4* has not yet been described. Future biochemical characterization of these proteins should reveal their molecular function.

Processing of the siRNA duplexes appears to start from the ends of both blunt-ended dsRNAs or dsRNAs with short (1-5 nt) 3' overhangs, and proceeds in approximately 21-23 nt steps. Long (~20 nt) 3' staggered ends on short dsRNAs suppress RNAi, possibly through interaction with single-stranded RNA-binding proteins. The suppression of RNAi by single-stranded regions flanking short dsRNA and the lack of siRNA formation from short 30 bp dsRNAs may explain why structured regions frequently encountered in mRNAs do not lead to activation of RNAi.

Without wishing to be bound by theory, we presume that the dsRNA-processing proteins or a subset of these remain associated with the siRNA duplex after the processing reaction. The orientation of the siRNA duplex relative to these proteins determines which of the two complementary strands functions in guiding target RNA degradation. Chemically synthesized siRNA duplexes guide cleavage of sense as well as antisense target RNA as they are able to associate with the protein components in either of the two possible orientation.

The remarkable finding that synthetic 21 and 22 nt siRNA duplexes can be used for efficient mRNA degradation provides new tools for sequence-specific regulation of gene expression in functional genomics as well as biomedical studies. The siRNAs may be effective in mammalian systems where long dsRNAs cannot be used due to the activation of the PKR response (Clemens, 1997). As such, the siRNA duplexes represent a new alternative to antisense or ribozyme therapeutics.

## **Example 2. RNA Interference in Human Tissue Cultures**

### **2.1. Methods**

#### **2.1.1. RNA preparation**

21 nt RNAs were chemically synthesized using Expedite RNA phosphoramidites and thymidine phosphoramidite (Proligo, Germany). Synthetic oligonucleotides were deprotected and gel-purified (Example 1), followed by Sep-Pak C18 cartridge (Waters, Milford, MA, USA) purification (Tuschl, 1993). The siRNA sequences targeting GL2 (Acc. X65324) and GL3 luciferase (Acc. U47296) corresponded to the coding regions 153-173 relative to the first nucleotide of the start codon; siRNAs targeting RL (Acc. AF025846) corresponded to region 119-129 after the start codon. Longer RNAs were transcribed with T7 RNA polymerase from PCR products, followed by gel and Sep-Pak purification. The 49 and 484 bp GL2 or GL3 dsRNAs corresponded to position 113-161 and 113-596, respectively, relative to the start of translation; the 50 and 501 bp RL dsRNAs corresponded to positions 118-167 and 118-618, respectively. PCR templates for dsRNA synthesis targeting humanized GFP (hG) were amplified from pAD3 (Kehlenbach, 1998), whereby 50 and 501 bp hG dsRNA corresponded to position 118-167 and 118-618, respectively, to the start codon.

For annealing of siRNAs, 20  $\mu$ M single strands were incubated in annealing buffer (100 mM potassium acetate, 30 mM HEPES-KOH at pH 7.4, 2 mM magnesium acetate) for 1 min at 90°C followed by 1 h at 37°C. The 37°C incubation step was extended overnight for the 50 and 500 bp dsRNAs and these annealing reactions were performed at 8.4  $\mu$ M and  
5 0.84  $\mu$ M strand concentrations, respectively.

### 2.1.2. Cell Culture

S2 cells were propagated in Schneider's *Drosophila* medium (Life Technologies) supplemented with 10% FBS, 100 units/ml penicillin and 100  $\mu$ g/ml streptomycin at 25°C. 293, NIH/3T3, HeLa S3, and COS-7 cells were grown at 37°C in  
10 Dulbecco's modified Eagle's medium supplemented with 10% FBS, 100 units/ml penicillin and 100  $\mu$ g/ml streptomycin. Cells were regularly passaged to maintain exponential growth 24 h before transfection at approx. 80% confluency, mammalian cells were trypsinized and diluted 1:5 with fresh medium without antibiotics ( $1-3 \times 10^5$  cells/ml) and transferred to 24-  
15 well plates (500  $\mu$ l/well). S2 cells were not trypsinized before splitting. Transfection was carried out with Lipofectamine 2000 reagent (Life Technologies) as described by the manufacturer for adherent cell lines. Per well, 1.0  $\mu$ g pGL2-Control (Promega) or pGL3-Control (Promega), 0.1  $\mu$ g pRL-TK (Promega) and 0.28  $\mu$ g siRNA duplex or dsRNA, formulated into liposomes, were applied; the final volume was 600  $\mu$ l per well. Cells were  
20 incubated 20 h after transfection and appeared healthy thereafter. Luciferase expression was subsequently monitored with the Dual luciferase assay (Promega). Transfection efficiencies were determined by fluorescence microscopy for mammalian cell lines after co-transfection of 1.1  $\mu$ g hGFP-encoding pAD3 and 0.28  $\mu$ g invGL2 inGL2 siRNA and were 70-90%. Reporter plasmids were amplified in XL-1 Blue (Stratagene) and purified using the Qiagen  
25 EndoFree Maxi Plasmid Kit.

### 2.2. Results and Discussion

To test whether siRNAs are also capable of mediating RNAi in tissue culture, we synthesized 21 nt siRNA duplexes with symmetric 2 nt 3' overhangs directed against reporter  
30 genes coding for sea pansy (*Renilla reniformis*) and two sequence variants of firefly (*Photinus pyralis*, GL2 and GL3) luciferases (Fig. 8a, b). The siRNA duplexes were co-transfected

with the reporter plasmid combinations pGL2/pRL or pGL3/pRL into *D. melanogaster* Schneider (S2) cells or mammalian cells using cationic liposomes. Luciferase activities were determined 20 h after transfection. In all cell lines tested, we observed specific reduction of the expression of the reporter genes in the presence of cognate siRNA duplexes (Fig. 9a-j).

5 Remarkably, the absolute luciferase expression levels were unaffected by non-cognate siRNAs, indicating the absence of harmful side effects by 21 nt RNA duplexes (e.g., Fig. 10a-d for HeLa cells). In *D. melanogaster* S2 cells (Fig. 9a, b), the specific inhibition of luciferases was complete. In mammalian cells, where the reporter genes were 50- to 100-fold more strongly expressed, the specific suppression was less complete (Fig. 9c-j). GL2  
10 expression was reduced 3- to 12-fold, GL3 expression 9- to 25-fold and RL expression 1- to 3-fold, in response to the cognate siRNAs. For 293 cells, targeting of RL luciferase by RL siRNAs was ineffective, although GL2 and GL3 targets responded specifically (Fig. 9i, j). The lack of reduction of RL expression in 293 cells may be due to its 5- to 20-fold higher expression compared to any other mammalian cell line tested and/or to limited accessibility  
15 of the target sequence due to RNA secondary structure or associated proteins. Nevertheless, specific targeting of GL2 and GL3 luciferase by the cognate siRNA duplexes indicated that RNAi is also functioning in 293 cells.

The 2 nt 3' overhang in all siRNA duplexes, except for uGL2, was composed of (2' deoxy) thymidine. Substitution of uridine by thymidine in the 3' overhang was well tolerated  
20 in the *D. melanogaster in vitro* system and the sequence of the overhang was not critical for target recognition. The thymidine overhang was chosen, because it is supposed to enhance nuclease resistance of siRNAs in the tissue culture medium and within transfected cells. Indeed, the thymidine-modified GL2 siRNA was slightly more potent than the unmodified uGL2 siRNA in all cell lines tested (Fig. 9a, c, e, g, i). It is conceivable that further  
25 modifications of the 3' overhanging nucleotides may provide additional benefits to the delivery and stability of siRNA duplexes.

In co-transfection experiments, 25 nM siRNA duplexes with respect to the final volume of tissue culture medium were used (Fig. 9, 10). Increasing the siRNA concentration to 100 nM did not enhance the specific silencing effect but started to affect transfection  
30 efficiencies due to competition for liposome encapsulation between plasmid DNA and siRNA (data not shown). Decreasing the siRNA concentration to 1.5 nM did not reduce the

specific silencing effect (data not shown), even though the siRNAs were now only 2- to 20-fold more concentrated than the DNA plasmids. This indicates that siRNAs are extraordinarily powerful reagents for mediating gene silencing and that siRNAs are effective at concentrations several orders of magnitude below the concentrations applied in conventional antisense or ribozyme gene targeting experiments.

In order to monitor the effect of longer dsRNAs on mammalian cells, 50 and 500 bp dsRNAs cognate to the reporter genes were prepared. As non-specific control, dsRNAs from humanized GFP (hG)(Kehlenbach, 1998) was used. When dsRNAs were co-transfected in identical amounts (not concentrations) to the siRNA duplexes, the reporter gene expression was strongly and unspecifically reduced. This effect is illustrated for HeLa cells as a representative example (Fig. 10a-d). The absolute luciferase activities were decreased unspecifically 10- to 20-fold by 50 bp dsRNA and 20- to 200-fold by 500 bp dsRNA co-transfection, respectively. Similar unspecific effects were observed for COS-7 and NIH/3T3 cells. For 293 cells, a 10- to 20-fold unspecific reduction was observed only for 500 bp dsRNAs. Unspecific reduction in reporter gene expression by dsRNA >30 bp was expected as part of the interferon response.

Surprisingly, despite the strong unspecific decrease in reporter gene expression, we reproducibly detected additional sequence-specific, dsRNA-mediated silencing. The specific silencing effects, however, were only apparent when the relative reporter gene activities were normalized to the hG dsRNA controls (Fig. 10e, f). A 2- to 10-fold specific reduction in response to cognate dsRNA was observed, also in the other three mammalian cell lines tested (data not shown). Specific silencing effects with dsRNAs (356-1 662 bp) were previously reported in CHO-K1 cells, but the amounts of dsRNA required to detect a 2-to 4-fold specific reduction were about 20-fold higher than in our experiments (Ui-Tei, 2000). Also, CHO-K1 cells appear to be deficient in the interferon response. In another report, 293, NIH/3T3 and BHK-21 cells were tested for RNAi using luciferase/lacZ reporter combinations and 829 bp specific lacZ or 717 bp unspecific GFP dsRNA (Caplen, 2000). The failure of detecting RNAi in this case may be due to the less sensitive luciferase/lacZ reporter assay and the length differences of target and control dsRNA. Taken together, our results indicate that RNAi is active in mammalian cells, but that the silencing effect is difficult to detect if the interferon system is activated by dsRNA >30 bp.

In summary, we have demonstrated for the first time siRNA-mediated gene silencing in mammalian cells. The use of short siRNAs holds great promise for inactivation of gene function in human tissue culture and the development of gene-specific therapeutics.

### Example 3. Specific Inhibition of Gene Expression by RNA Interference

#### 3.1. Materials and Methods

##### 3.1.1. RNA preparation and RNAi assay

Chemical RNA synthesis, annealing, and luciferase-based RNAi assays were performed as described in Examples 1 or 2, or in previous publications (Tuschl *et al.*, 1999; Zamore *et al.*, 2000). All siRNA duplexes were directed against firefly luciferase, and the luciferase mRNA sequence was derived from pGEM-luc (GenBank acc. X65316) as described (Tuschl *et al.*, 1999). The siRNA duplexes were incubated in *D. melanogaster* RNAi/translation reaction for 15 min prior to addition of mRNAs. Translation-based RNAi assays were performed at least in triplicates.

For mapping of sense target RNA cleavage, a 177-nt transcript was generated, corresponding to the firefly luciferase sequence between positions 113-273 relative to the start codon, followed by the 17-nt complement of the SP6 promoter sequence. For mapping of antisense target RNA cleavage, a 166-nt transcript was produced from a template, which was amplified from plasmid sequence by PCR using 5' primer

TAATACGACTCACTATAGAGCCCATATCGTTTCATA (T7, promoter underlined) and 3' primer AGAGGATGGAACCGCTGG. The target sequence corresponds to the complement of the firefly luciferase sequence between positions 50-215 relative to the start codon. Guanylyl transferase labeling was performed as previously described (Zamore *et al.*, 2000). For mapping of target RNA cleavage, 100 nM siRNA duplex was incubated with 5 to 10 nM target RNA in *D. melanogaster* embryo lysate under standard conditions (Zamore *et al.*, 2000) for 2 h at 25°C. The reaction was stopped by the addition of 8 volumes of proteinase K buffer (200 mM Tris-HCl pH 7.5, 25 mM EDTA, 300 mM NaCl, 2% w/v sodium dodecyl sulfate). Proteinase K (E.M. Merck, dissolved in water) was added to a final concentration of 0.6 mg/ml. The reactions were then incubated for 15 min at 65°C, extracted with phenol/chloroform/isoamyl alcohol (25:24:1) and precipitated with 3 volumes of ethanol. Samples were located on 6% sequencing gels. Length standards were generated by

partial RNase T1 digestion and partial base hydrolysis of the cap labeled sense or antisense target RNAs.

### 3.2. Results

#### 3.2.1. Variation of the 3' overhang in duplexes of 21-nt siRNAs

As described above, 2 or 3 unpaired nucleotides at the 3' end of siRNA duplexes were more efficient in target RNA degradation than the respective blunt-ended duplexes. To perform a more comprehensive analysis of the function of the terminal nucleotides, we synthesized five 21-nt sense siRNAs, each displayed by one nucleotide relative to the target RNA, and eight 21-nt antisense siRNAs, each displaced by one nucleotide relative to the target (FIG. 11A). By combining sense and antisense siRNAs, eight series of siRNA duplexes with synthetic overhanging ends were generated covering a range of 7-nt 3' overhang to 4-nt 5' overhang. The interference of siRNA duplexes was measured using the dual luciferase assay system (Tuschl *et al.*, 1999; Zamore *et al.*, 2000). siRNA duplexes were directed against firefly luciferase mRNA, and sea pansy luciferase mRNA was used as an internal control. The luminescence ratio of target to control luciferase activity was determined in the presence of siRNA duplex and was normalized to the ratio observed in the absence of dsRNA. For comparison, the interference ratios of long dsRNAs (39 to 504 bp) are shown in FIG. 11B. The interference ratios were determined at concentrations of 5 nM for long dsRNAs (FIG. 11A) and at 100 nM for siRNA duplexes (FIG. 11 C-J). The 100 nM concentrations of siRNAs was chosen, because complete processing of 5 nM 504 bp dsRNA would result in 120 nM total siRNA duplexes.

The ability of 21-nt siRNA duplexes to mediate RNAi is dependent on the number of overhanging nucleotides or base pairs formed. Duplexes with four to six 3' overhanging nucleotides were unable to mediate RNAi (FIG. 11C-F), as were duplexes with two or more 5' overhanging nucleotides (FIG. 11G-J). The duplexes with 2-nt 3' overhangs were most efficient in mediating RNA interference, though the efficiency of silencing was also sequence-dependent, and up to 12-fold differences were observed for different siRNA duplexes with 2-nt 3' overhangs (compare FIG. 11D-H). Duplexes with blunted ends, 1-nt 5' overhang or 1- to 3-nt 3' overhangs were sometimes functional. The small silencing effect observed for the siRNA duplex with 7-nt 3' overhang (FIG. 11C) may be due to an antisense



effect of the long 3' overhang rather than due to RNAi. Comparison of the efficiency of RNAi between long dsRNAs (Fig. 11B) and the most effective 21-nt siRNA duplexes (Fig. 11E, G, and H) indicates that a single siRNA duplex at 100 nM concentration can be as effective as 5 nM 504 bp dsRNA.

5

### 3.2.2. Length variation of the sense siRNA paired to an invariant 21-nt antisense siRNA

In order to investigate the effect of length of siRNA on RNAi, we prepared three series of siRNA duplexes, combining three 21-nt antisense strands with eight, 18- to 25-nt sense strands. The 3' overhang of the antisense siRNA was fixed to 1, 2, or 3 nt in each siRNA duplex series, while the sense siRNA was varied at its 3' end (FIG. 12A). Independent of the length of the sense siRNA, we found that duplexes with 2-nt 3' overhang of antisense siRNA (FIG. 12C) were more active than those with 1- or 3-nt 3' overhang (FIG. 12 B, D). In the first series, with 1-nt 3' overhang of antisense siRNA, duplexes with a 21- and 22-nt sense siRNAs, carrying a 1- and 2-nt 3' overhang of sense siRNA, respectively, were most active. Duplexes with 19- to 25-nt sense siRNAs were also able to mediate RNAi, but to a lesser extent. Similarly, in the second series, with 2-nt overhang of antisense siRNA, the 21-nt siRNA duplex with 2-nt 3' overhang was most active, and any other combination with the 18- to 25-nt sense siRNAs was active to a significant degree. In the last series, with 3-nt antisense siRNA 3' overhang, only the duplex with a 20-nt sense siRNA and the 2-nt sense 3' overhang was able to reduce target RNA expression. Together, these results indicate that the length of the siRNA as well as the length of the 3' overhang are important, and that duplexes of 21-nt siRNAs with 2-nt 3' overhang are optimal for RNAi.

### 3.2.3 Length variation of siRNA duplexes with a constant 2-nt 3' overhang

We next examined the effect of simultaneously changing the length of both siRNA strands by maintaining symmetric 2-nt 3' overhangs (FIG. 13A). Two series of siRNA duplexes were prepared including the 21-nt siRNA duplex of FIG. 11H as reference. The length of the duplexes was varied between 20 to 25 bp by extending the base-paired segment at the 3' end of the sense siRNA (FIG. 13B) or at the 3' end of the antisense siRNA (FIG. 13C). Duplexes of 20 to 23 bp caused specific repression of target luciferase activity, but the

30

21-nt siRNA duplex was at least 8-fold more efficient than any of the other duplexes. 24- and 25-nt siRNA duplexes did not result in any detectable interference. Sequence-specific effects were minor as variations on both ends of the duplex produced similar effects.

#### 5           3.2.4. 2'-Deoxy and 2'-O-methyl-modified siRNA duplexes

To assess the importance of the siRNA ribose residues for RNAi duplexes with 21-nt siRNAs and 2-nt 3' overhangs with 2' deoxy- or 2'-O-methyl-modified strands were examined (FIG. 14). Substitution of 2-nt 3' overhangs by 2'-deoxy nucleotides had no effect, and even the replacement of two additional ribonucleotides adjacent to the overhangs in the paired  
10 region produced significantly active siRNAs. Thus, 8 out of 42 nt of an siRNA duplex were replaced by DNA residues without loss of activity. Complete substitution of one or both siRNA strands by 2'-deoxy residues, however, abolished RNAi, as did substitution by 2'-O-methylresidues.

#### 15           3.2.5. Definition of target RNA cleavage sites

Target RNA cleavage positions were previously determined for 22-nt siRNA duplexes and for a 21-nt/22-nt duplex. It was found that the position of the target RNA cleavage was located in the center of the region covered by the siRNA duplex, 11 or 12 nt downstream of the first nucleotide that was complementary to the 21- or 22-nt siRNA guide  
20 sequence. Five distinct 21-nt siRNA duplexes with 2-nt 3' overhang (FIG. 15A) were incubated with 5' cap-labeled sense or antisense target RNA in *D. melanogaster* lysate (Tuschl *et al.*, 1999; Zamore *et al.*, 2000). The 5' cleavage products were resolved on sequencing gels (FIG. 15B). The amount of sense target RNA cleaved correlates with the efficiency of siRNA duplexes determined in the translation-based assay, and siRNA duplexes  
25 1, 2 and 4 (FIG. 15B and 11 H, G, E) cleave target RNA faster than duplexes 3 and 5 (FIG. 15B and 11 F, D). Notably, the sum of radioactivity of the 5' cleavage product and the input target RNA were not constant over time, and the 5' cleavage products did not accumulate. Presumably, the cleavage products, once released from the siRNA-endonuclease complex, are rapidly degraded due to the lack of either of the poly(A) tail or the 5'-cap.

30           The cleavage sites for both sense and antisense target RNAs were located in the middle of the region spanned by the siRNA duplexes. The cleavage sites for each target

produced by the five different duplexes varied by one nucleotide according to the 1-nt displacement of the duplexes along the target sequences. The targets were cleaved precisely 11 nt downstream of the target position complementary to the 3'-most nucleotide of the sequence-complementary guide siRNA (FIG. 15A, B).

5 In order to determine, whether the 5' or the 3' end of the guide siRNA sets the ruler for target RNA cleavage, we devised the experimental strategy outlined in FIG. 16A and B. A 21-nt antisense siRNA, which was kept invariant for this study, was paired with sense siRNAs that were modified at either of their 5' or 3' ends. The position of sense and antisense target RNA cleavage was determined as described above. Changes in the 3' end of the sense siRNA, monitored for 1-nt 5' overhang to 6-nt 3' overhang, did neither effect the position of sense nor antisense target RNA cleavage (FIG. 16C). Changes in the 5' end of the sense siRNA did not affect the sense target RNA cleavage (FIG. 16D, top panel), which was expected because the antisense siRNA was unchanged. However, the antisense target RNA cleavage was affected and strongly dependent on the 5' end of the sense siRNA (FIG. 16D, bottom panel). The antisense target was only cleaved, when the sense siRNA was 20 or 21 nt in size, and the position of cleavage different by 1-nt, suggesting that the 5' end of the target-recognizing siRNA sets the ruler for target RNA cleavage. The position is located between nucleotide 10 and 11 when counting in upstream direction from the target nucleotide paired to the 5'-most nucleotide of the guide siRNA (see also FIG. 15A).

### 3.2.6. Sequence effects and 2'-deoxy substitutions in the 3' overhang

20 A 2-nt 3' overhang is preferred for siRNA function. We wanted to know, if the sequence of the overhanging nucleotides contributes to target recognition, or if it is only a feature required for reconstitution of the endonuclease complex (RISC or siRNP). We synthesized sense and antisense siRNAs with AA, CC, GG, UU, and UG 3' overhangs and included the 2'-deoxy modifications TdG and TT. The wild-type siRNAs contained AA in the sense 3' overhang and UG in the antisense 3' overhang (AA/UG). All siRNA duplexes were functional in the interference assay and reduced target expression at least 5-fold (FIG. 17). The most efficient siRNA duplexes that reduced target expression more than 10-fold, were of the sequence type NN/UG, NN/UU, NN/TdG, and NN/TT (N, any nucleotide). siRNA duplexes with an antisense siRNA 3' overhang of AA, CC or GG were less active by

a factor 2 to 4 when compared to the wild-type sequence UG or the mutant UU. This reduction in RNAi efficiency is likely due to the contribution of the penultimate 3' nucleotide to sequence-specific target recognition, as the 3' terminal nucleotide was changed from G to U without effect.

Changes in the sequence of the 3' overhang of the sense siRNA did not reveal any sequence-dependent effects, which was expected, because the sense siRNA must not contribute to sense target mRNA recognition.

### 3.2.7. Sequence specificity of target recognition

In order to examine the sequence-specificity of target recognition, we introduced sequence changes into the paired segments of siRNA duplexes and determined the efficiency of silencing. Sequence changes were introduced by inverting short segments of 3- or 4-nt length or as point mutations (FIG. 18). The sequence changes in one siRNA strand were compensated in the complementary siRNA strand to avoid perturbing the base-paired siRNA duplex structure. The sequence of all 2-nt 3' overhangs was TT (T, 2'-deoxythymidine) to reduce costs of synthesis. The TT/TT reference siRNA duplex was comparable in RNAi to the wild-type siRNA duplex AA/UG (FIG. 17). The ability to mediate reporter mRNA destruction was quantified using the translation-based luminescence assay. Duplexes of siRNAs with inverted sequence segments showed dramatically reduced ability for targeting the firefly luciferase reporter (FIG. 18). The sequence changes located between the 3' end and the middle of the antisense siRNA completely abolished target RNA recognition, but mutations near the 5' end of the antisense siRNA exhibited a small degree of silencing. Transversion of the A/U base pair located directly opposite of the predicted target RNA cleavage site, or one nucleotide further away from the predicted site, prevented target RNA cleavage, therefore indicating that a single mutation within the center of an siRNA duplex discriminated between mismatched targets.

### 3.3. Discussion

siRNAs are valuable reagents for inactivation of gene expression, not only in insect cells, but also in mammalian cells, with a great potential for therapeutic application. We have systematically analyzed the structural determinants of siRNA duplexes required to promote

efficient target RNA degradation in *D. melanogaster* embryo lysate, thus providing rules for the design of most potent siRNA duplexes. A perfect siRNA duplex is able to silence gene expression with an efficiency comparable to a 500 bp dsRNA, given that comparable quantities of total RNA are used.

5

### 3.4. The siRNA user guide

Efficiently silencing siRNA duplexes are preferably composed of 21-nt antisense siRNAs, and should be selected to form a 19 bp double helix with 2-nt 3' overhanging ends. 2'-deoxy substitutions of the 2-nt 3' overhanging ribonucleotides do not affect RNAi, but help to reduce the costs of RNA synthesis. More extensive 2'-deoxy or 2'-O-methylmodifications, however, reduce the ability of siRNAs to mediate RNAi, probably by interfering with protein association for siRNP assembly.

Target recognition is a highly sequence-specific process, mediated by the siRNA complementary to the target. The 3'-most nucleotide of the guide siRNA does not contribute to specificity of target recognition, while the penultimate nucleotide of the 3' overhang affects target RNA cleavage, and a mismatch reduces RNAi 2- to 4-fold. The 5' end of a guide siRNA also appears more permissive for mismatched target RNA recognition when compared to the 3' end. Nucleotides in the center of the siRNA, located opposite the target RNA cleavage site, are important specificity determinants and even single nucleotide changes reduce RNAi to an undetectable level. This suggests that siRNA duplexes may be able to discriminate mutant or polymorphic alleles in gene targeting experiments, which may become an important feature for future therapeutic developments.

Sense and antisense siRNAs, when associated with the protein components of the endonuclease complex or its commitment complex, were suggested to play distinct roles; the relative orientation of the siRNA duplex in this complex defines which strand can be used for target recognition. Synthetic siRNA duplexes have dyad symmetry with respect to the double-helical structure, but not with respect to sequence. The association of siRNA duplexes with the RNAi proteins in the *D. melanogaster* lysate will lead to formation of two asymmetric complexes. In such hypothetical complexes, the chiral environment is distinct for sense and antisense siRNA, hence their function. The prediction obviously does not apply to palindromic siRNA sequences, or to RNAi proteins that could associate as

homodimers. To minimize sequence effects, which may affect the ratio of sense and antisense-targeting siRNPs, we suggest to use siRNA sequences with identical 3' overhanging sequences. We recommend adjusting the sequence of the overhang of the sense siRNA to that of the antisense 3' overhang, because the sense siRNA does not have a target in typical knockdown experiments. Asymmetry in reconstitution of sense and antisense-cleaving siRNPs could be (partially) responsible for the variation in RNAi efficiency observed for various 21-nt siRNA duplexes with 2-nt 3' overhangs used in this study (FIG. 14). Alternatively, the nucleotide sequence at the target site and/or the accessibility of the target RNA structure may be responsible for the variation in efficiency for these siRNA duplexes.

A number of embodiments of the invention have been described. Nevertheless, it will be understood that various modifications may be made without departing from the spirit and scope of the invention. Accordingly, other embodiments are within the scope of the following claims.

#### References

Bass, B. L. (2000). Double stranded RNA as a template for gene silencing. *Cell* 101, 235-238.

Bosher, J. M., and Labouesse, M. (2000). RNA interference: genetic wand and genetic watchdog. *Nat. Cell Biol.* 2, E3 1-36.

Caplen, N. J., Fleenor, J., Fire, A., and Morgan, R. A. (2000). dsRNA-mediated gene silencing in cultured *Drosophila* cells: a tissue culture model for the analysis of RNA interference. *Gene* 252, 95-105.

Catalanotto, C., Azzalin, G., Macino, G., and Cogoni, C. (2000). Gene silencing in worms and fungi. *Nature* 404, 245.

Chanfreau, G., Buckle, M., and Jacquier, A. (2000). Recognition of a conserved class of RNA tetraloops by *Saccharomyces cerevisiae* RNase III. *Proc. Natl. Acad. Sci. USA* 97, 3142-3147.

Clemens, M. J. (1997). PKR-a protein kinase regulated by double stranded RNA. *Int. J. Biochem. Cell Biol.* 29, 945-949.

Cogoni, C., and Macino, G. (1999). Homology-dependent gene silencing in plants and fungi: a number of variations on the same theme. *Curr. Opin. Microbiol.* 2, 657-662.

Dalmay, T., Hamilton, A., Rudd, S., Angell, S., and Baulcombe, D. C. (2000). An RNA dependent RNA polymerase gene in *Arabidopsis* is required for posttranscriptional gene silencing mediated by a transgene but not by a virus. *Cell* 101, 543-553.

Demburg, A. F., Zalevsky, J., Colaiacovo, M. P., and Villeneuve, A. M. (2000). Transgene-mediated cosuppression in the *C. elegans* germ line. *Genes & Dev.* 14, 1578-1583.

Dunn, J. J. (1982). Ribonuclease III. In *The enzymes*, vol. 15, part B, P. D. Boyer, ed. (New York: Academic Press), pp. 485-499.

Filippov, V., Solovyev, V., Filippova, M., and Gill, S. S. (2000). A novel type of RNase III family proteins in eukaryotes. *Gene* 245, 213-221.

Fire, A. (1999). RNA-triggered gene silencing. *Trends Genet.* 15, 358-363.

Fire, A., Xu, S., Montgomery, M. K., Kostas, S. A., Driver, S. E., and Mello, C. C. (1998). Potent and specific genetic interference by double stranded RNA in *Caenorhabditis elegans*. *Nature* 391, 806-811.

Grishok, A., Tabara, H., and Mello, C. C. (2000). Genetic requirements for inheritance of RNAi in *C. elegans*. *Science* 287, 2494-2497.

Hamilton, A. J., and Baulcombe, D. C. (1999). A species of small antisense RNA in posttranscriptional gene silencing in plants. *Science* 286, 950-952.

Hammond, S. M., Bernstein, E., Beach, D., and Hannon, G. J. (2000). An RNA-directed nuclease mediates post-transcriptional gene silencing in *Drosophila* cells. *Nature* 404, 293-296.

Jacobsen, S. E., Running, M. P., and M., M. E. (1999). Disruption of an RNA helicase/RNase III gene in *Arabidopsis* causes unregulated cell division in floral meristems. *Development* 126, 5231-5243.

Jensen, S., Gassama, M. P., and Heidmann, T. (1999). Taming of transposable elements by homology-dependent gene silencing. *Nat. Genet.* 21, 209-212.

Kehlenbach, R. H., Dickmanns, A. & Gerace, L. (1998). Nucleocytoplasmic shuttling factors including Ran and CRM1 mediate nuclear export of NFAT *In vitro*. *J. Cell Biol.* 141, 863-874.

Kennerdell, J. R., and Carthew, R. W. (1998). Use of dsRNA-mediated genetic interference to demonstrate that *frizzled* and *frizzled 2* act in the wingless pathway. *Cell* 95, 1017-1026.

Ketting, R. F., Haverkamp, T. H., van Lucen, H. G., and Plasterk, R. H. (1999). Mut-7 of *C. elegans*, required for transposon silencing and RNA interference, is a homolog of Werner syndrome helicase and Rnase D. *Cell* 99, 133-141.

Ketting, R. F., and Plasterk, R. H. (2000). A genetic link between co-suppression and RNA interference in *C. elegans*. *Nature* 404, 296-298.

Lucy, A. P., Guo, H. S., Li, W. X., and Ding, S. W. (2000). Suppression of post transcriptional gene silencing by a plant viral protein localized in the nucleus. *EMBO J.* 19, 1672-1680.

Matsuda, S., Ichigotani, Y., Okuda, T., Irimura, T., Nakatsugawa, S., and Hamaguchi, M. (2000). Molecular cloning and characterization of a novel human gene (HERNA) which encodes a putative RNA-helicase. *Biochim. Biophys. Acta* 31, 1-2.

Milligan, J. F., and Uhlenbeck, O. C. (1989). Synthesis of small RNAs using T7 RNA polymerase. *Methods Enzymol.* 180, 51-62.

Mourrain, P., Beclin, C., Elmayan, T., Feuerbach, F., Godon, C., Morel, J. B., Jouette, D., Lacombe, A. M., Nikic, S., Picaud, N., Remoue, K., Sanial, M., Vo, T. A., and Vaucheret, H. (2000). Arabidopsis SGS2 and SGS3 genes are required for posttranscriptional gene silencing and natural virus resistance. *Cell* 101, 533-542.

Ngo, H., Tschudi, C., Gull, K., and Ullu, B. (1998). Double stranded RNA induces mRNA degradation in *Trypanosoma brucei*. *Proc. Natl. Acad. Sci. USA* 95, 14687-14692.

Nicholson, A. W. (1999). Function, mechanism and regulation of bacterial ribonucleases. *FEMS Microbiol. Rev.* 23, 371-390.

Oelgeschlager, M., Lanam, J., Geissert, D., and De Robertis, E. M. (2000). The evolutionarily conserved BMP-binding protein Twisted gastrulation promotes BMP signalling. *Nature* 405, 757-763.

Pan, T., and Uhlenbeck, O. C. (1992). *In vitro* selection of RNAs that undergo autolytic cleavage with Pb<sup>2+</sup>. *Biochemistry* 31, 3887-3895.

Pelissier, T., and Wassenegger, M. (2000). A DNA target of 30 bp is sufficient for RNA directed methylation. *RNA* 6, 55-65.



Plasterk, R. H., and Ketting, R. F. (2000). The silence of the genes. *Curr. Opin. Genet. Dev.* 10, 562-567.

Ratcliff, F. G., MacFarlane, S. A., and Baulcombe, D. C. (1999). Gene Silencing without DNA. RNA-mediated cross-protection between viruses. *Plant Cell* 11, 1207-1216.

5 Robertson, H. D. (1990). *Escherichia coli* ribonuclease III. *Methods Enzymol.* 181, 189-202.

Robertson, H. D. (1982). *Escherichia coli* ribonuclease III cleavage sites. *Cell* 30,669-672.

10 Romaniuk, E., McLaughlin, L. W., Neilson, T., and Romaniuk, P. J. (1982). The effect of acceptor oligoribonucleotide sequence on the T4 RNA ligase reaction. *Eur J Biochem* 125, 639-643.

Sharp, P. A. (1999). RNAi and double-strand RNA. *Genes & Dev.* 13, 139-141.

Sijen, T., and Kooter, J. M. (2000). Post-transcriptional gene-silencing: RNAs on the attack or on the defense? *Bioessays* 22, 520-531.

15 Smardon, A., Spoerke, J., Stacey, S., Klein, M., Mackin, N., and Maine, E. (2000). EGO-1 is related to RNA-directed RNA polymerase and functions in germ-line development and RNA interference in *C. elegans*. *Curr. Biol.* 10, 169-178.

20 Svoboda, P., Stein, P., Hayashi, H., and Schultz, R. M. (2000). Selective reduction of dormant maternal mRNAs in mouse oocytes by RNA interference. *Development* 127, 4147-4156.

Tabara, H., Sarkissian, M., Kelly, W. G., Fleenor, J., Grishok, A., Timmons, L., Fire, A., and Mello, C. C. (1999). The rde-1 gene, RNA interference, and transposon silencing in *C. elegans*. *Cell* 99, 123-132.

25 Tuschl, T., Ng, M. M., Pieken, W., Benseler, F., and Eckstein, F. (1993). Importance of exocyclic base functional groups of central core guanosines for hammerhead ribozyme activity. *Biochemistry* 32,11658-11668.

Tuschl, T., Sharp, P. A., and Bartel, D. P. (1998). Selection *in vitro* of novel ribozymes from a partially randomized U2 and U6 snRNA library. *EMBO J.* 17, 2637-2650.

30 Tuschl, T., Zamore, P. D., Lehmann, R., Bartel, D. P., and Sharp, P. A. (1999). Targeted mRNA degradation by double stranded RNA *in vitro*. *Genes & Dev.* 13, 3191-3197.

Ui-Tei, K., Zenno, S., Miyata, Y. and Saigo, K. (2000). Sensitive assay of RNA interference in *Drosophila* and Chinese hamster cultured cells using firefly luciferase gene as target. *FEBS Letters* 479, 79-82.

5 Verma, S., and Eckstein, F. (1999). Modified oligonucleotides: Synthesis and strategy for users. *Annu. Rev. Biochem.* 67, 99-134.

Voinnet, O., Lederer, C., and Baulcombe, D. C. (2000). A viral movement protein prevents spread of the gene silencing signal in *Nicotiana benthamiana*. *Cell* 103, 157-167.

Wassenegger, M. (2000). RNA-directed DNA methylation. *Plant Mol. Biol.* 43, 203-220.

10 Wianny, F., and Zernicka-Goetz, M. (2000). Specific interference with gene function by double stranded RNA in early mouse development. *Nat. Cell Biol.* 2, 70-75.

Wu, H., Xu, H., Miraglia, L. J., and Crooke, S. T. (2000). Human RNase III is a 160 kD Protein Involved in Preribosomal RNA Processing. *J. Biol. Chem.* 17, 17.

15 Yang, D., Lu, H. and Erickson, J. W. (2000). Evidence that processed small dsRNAs may mediate sequence-specific mRNA degradation during RNAi in *drosophila* embryos. *Curr. Biol.*, 10, 1191-1200.

Zamore, P. D., Tuschl, T., Sharp, P. A., and Bartel, D. P. (2000). RNAi: Double stranded RNA directs the ATP-dependent cleavage of mRNA at 21 to 23 nucleotide intervals. *Cell* 101, 25-33.

20 Zhang, K., and Nicholson, A. W. (1997). Regulation of ribonuclease III processing by double-helical sequence antideterminants. *Proc. Natl. Acad. Sci. USA* 94, 13437-1 3441.

**WHAT IS CLAIMED IS:**

1. A method of treating a subject comprising:

- 5           (a) identifying a subject having, or at risk for having, unwanted cell proliferation;  
            (b) providing an siRNA that is homologous to a gene that can promote cell proliferation; and  
            (c) administering the siRNA to the subject,  
10           thereby treating the subject.

2. The method of claim 1, wherein the subject has a malignant or nonmalignant cell proliferation.

15           3. The method of claim 1, wherein the gene encodes one of the group consisting of a growth factor, growth factor receptor, kinase, adaptor protein, and transcription factor.

            4. The method of claim 1, wherein the gene is one of the group consisting of RAS, JNK, c-MYC, cyclin D, and beta-catenin.

20

            5. The method of claim 1, wherein the subject has, or is at risk for having, one of the group consisting lung cancer, breast cancer, colon cancer, or liver cancer.

6. A method of treating a subject comprising:

- 25           (a) identifying a subject having, or at risk for having, a disorder associated with a viral infection;  
            (b) providing an siRNA that is homologous to a gene that can mediate viral function; and  
            (c) administering the siRNA to the subject,  
30           thereby treating the subject.

7. The method of claim 6, wherein the gene is a viral gene or an endogenous gene of the subject.

8. The method of claim 6, wherein the subject has, or is at risk for having, Acquired Immune Deficiency Syndrome (AIDS), asthma, or Hodgkin's disease.

9. The method of claim 6, wherein the subject is infected by, or is at risk for being infected by a virus in the group consisting of Human Immunodeficiency Virus (HIV), a hepatitis virus, respiratory syncytial virus (RSV), or Epstein Barr Virus (EBV).

10. The method of claim 9, wherein the hepatitis virus is hepatitis A, B, or C.

11. The method of claim 9, wherein the gene is Gag or Rev gene of HIV.

12. A method of treating a subject comprising:

(a) identifying a subject having, or at risk for having, a disorder characteristic of an unwanted immune response;

(b) providing an siRNA that is homologous to a gene that can affect the unwanted immune response; and

(c) administering the siRNA to the subject, thereby treating the subject.

13. The method of claim 12, wherein the unwanted immune response is caused by an autoimmune disease or disorder.

14. The method of claim 12, wherein the gene is endogenous to the subject.

15. The method of claim 12, wherein the subject has, or is at risk for having, inflammatory bowel disease, asthma, multiple sclerosis, or a reperfusion injury.

16. The method of claim 15, wherein the reperfusion injury is associated with acute myocardial infarction, unstable angina, cardiopulmonary bypass, or angioplasty.

17. The method of claim 12, wherein the gene encodes a complement factor.

5

18. The method of claim 12, wherein the gene encodes a protein selected from the group consisting of VLA4, VCAM, ICAM, E-selectin (ELAM), C5 complement, TNF $\alpha$ , TNF $\beta$ , IL-2, IL-2R, IL-4, IL-4R, IL-5, CCR3, Gro1, Gro2, and Gro3.

10

19. A method of treating a subject comprising:

(a) identifying a subject having, or at risk for having, a disorder characterized by acute or chronic pain;

(b) providing an siRNA that is homologous to a gene that can affect the processing of pain; and

15

(c) administering the siRNA to the subject, thereby treating the subject.

20. The method of claim 19, wherein the gene encodes a component of an ion channel, a neurotransmitter receptor, or a ligand of a neurotransmitter receptor.

20

21. A method of treating a subject comprising:

(a) identifying a subject having, or at risk for having, a neurological disease or disorder;

(b) providing an siRNA that is homologous to a gene that can affect the neurological disease or disorder; and

25

(c) administering the siRNA to the subject, thereby treating the subject.

22. The method of claim 21, wherein the subject has, or is at risk for having, Alzheimer's disease, Parkinson's disease, or a trinucleotide repeat disorder.

30

23. The method of claim 22, wherein the trinucleotide repeat disorder is Huntington's disease.

24. The method of claim 21, wherein the gene encodes a protein selected from the group consisting of APP, PSEN1, PSEN2, and  $\alpha$ -synuclein.

25. A method of treating a subject having, or at risk for having, a loss of heterozygosity (LOH), the method comprising:

- (a) identifying a subject having, or at risk for having, LOH;
- (b) determining the genotype of an allele of a gene in a region of LOH in an LOH cell;
- (c) determining the genotype of both alleles of the gene in a normal cell;
- (d) providing an siRNA that is homologous to an allele of the gene in the LOH cell, but not homologous to an allele of the gene in a normal cell; and
- (e) administering the siRNA to the subject, thereby treating the subject.

26. The method of claim 25, wherein the subject has, or is at risk for having, a cancer.

27. The method of claim 25, wherein the siRNA is homologous to an allele having a single nucleotide polymorphism (SNP).

28. The method of claim 25, wherein the LOH cell is a tumor cell.

29. The method of claim 25, wherein the siRNA is homologous to an allele of the large subunit of human RNA polymerase II (POLR2A), the replication protein A 70-kD subunit, replication protein A 32-kD subunit, ribonucleotide reductase, thymidilate synthase, TATA associated factor 2H, ribosomal protein S14, eukaryotic initiation factor 5A, alanyl tRNA synthetase, cysteinyl tRNA synthetase, NaK ATPase, alpha-1 subunit, or transferrin receptor.

30. Use of double-stranded RNA molecules, wherein each RNA strand has a length of 19-25 nucleotides for the manufacture of an agent for mediating target-specific nucleic acid modifications in mammalian cells or organisms.
31. The use of claim 30, wherein at least one strand has a 3'-overhang from 1-5 nucleotides.
32. The use of claims 30 or 31, wherein the target-specific nucleic acid modification is RNA interference.
33. The use of any one of claims 30-32 for silencing at least one gene which promotes unwanted cell proliferation.
34. The use of claim 33 wherein the gene is selected from the group consisting of growth factor genes, growth factor receptor genes, adaptor protein genes, genes encoding a G protein superfamily molecule and genes encoding a transcription factor.
35. The use of any one of claims 30-32 for silencing at least one gene which mediates angiogenesis.
36. The use of any one of claims 30-32 for silencing at least one viral gene or cellular gene which mediates viral function.
37. The use of any one of claims 30-32 for silencing at least one gene from a bacterial, amoebic, parasitic or fungal pathogen.
38. The use of any one of claims 30-32 for silencing at least one gene which mediates an unwanted immune response.

-71-

39. The use of any one of claims 30-32 for silencing at least one gene which mediates the processing of pain.
40. The use of any one of claims 30-32 for silencing at least one gene which mediates a neurological disease or disorder.
41. The use of any one of claims 30-40 for the allele-specific silencing of at least one gene.
42. The use of any one of claims 30-41 wherein the agent comprises the double-stranded RNA molecule as an active ingredient.
43. The use of any one of claims 30-41 wherein the agent comprises vectors capable of expressing the double-stranded RNA molecule as an active ingredient.



FIGURE 1A

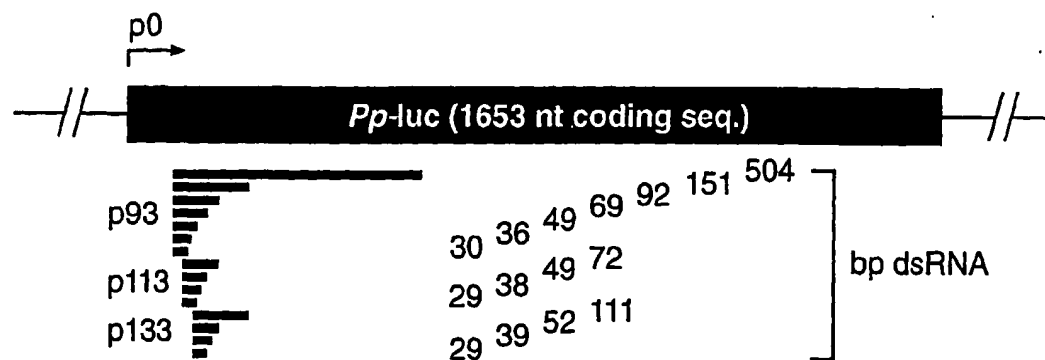
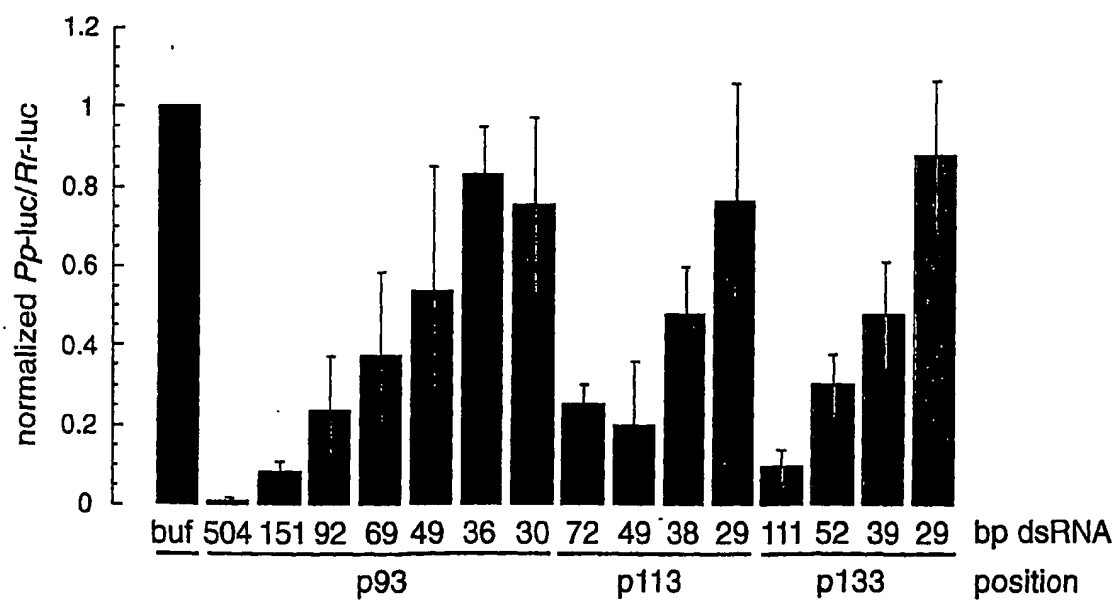


FIGURE 1B



2/23

FIGURE 2

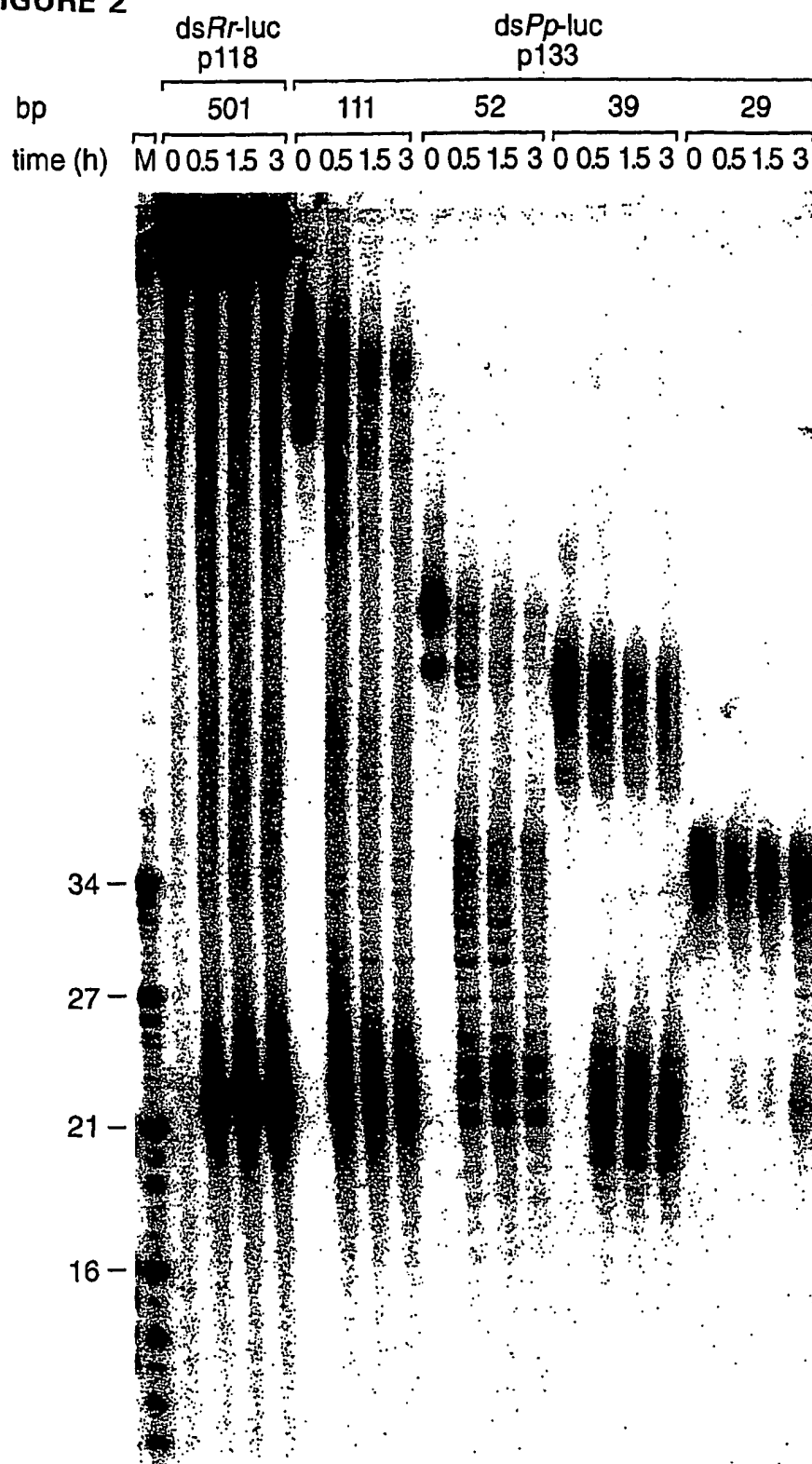
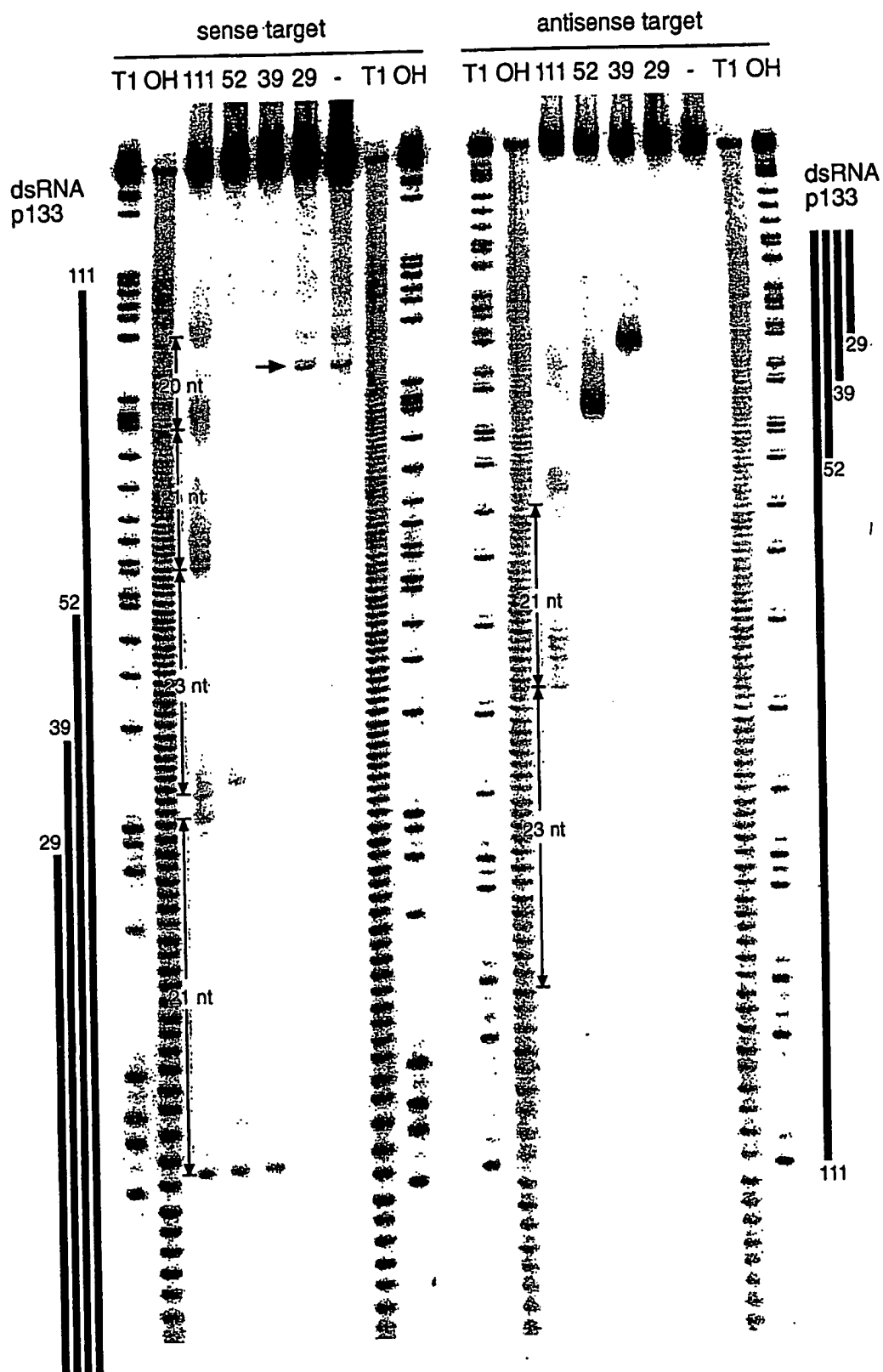


FIGURE 3A



4/23

FIGURE 3B

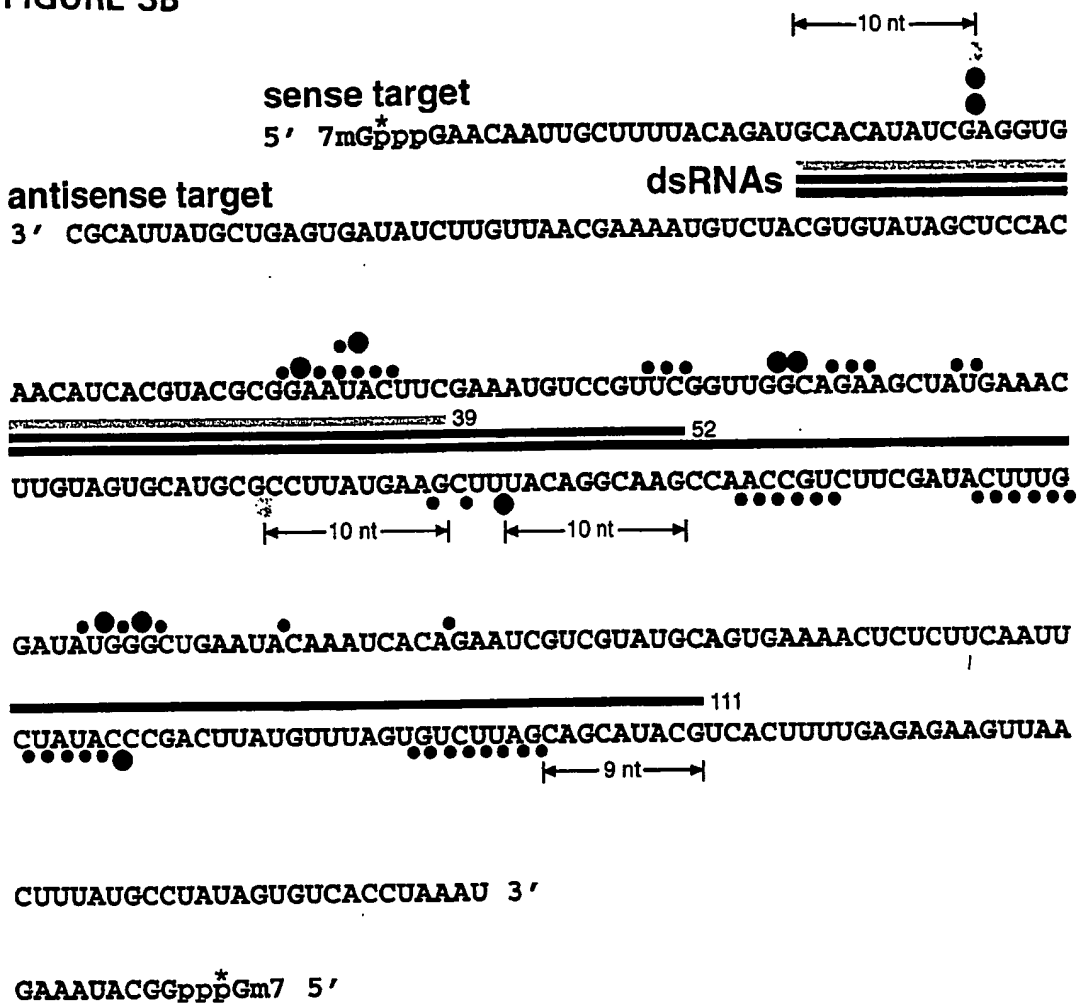
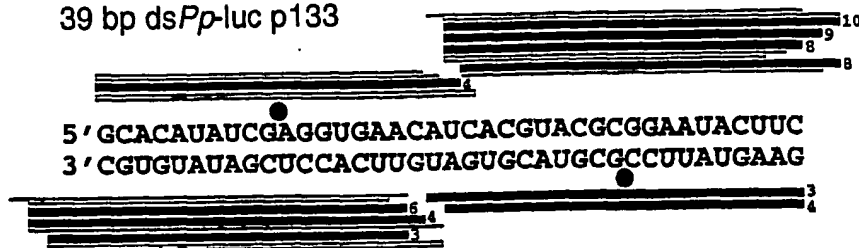
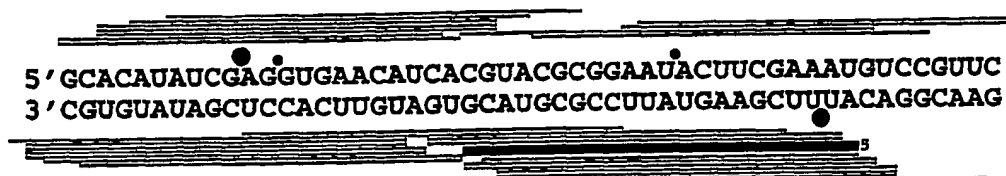
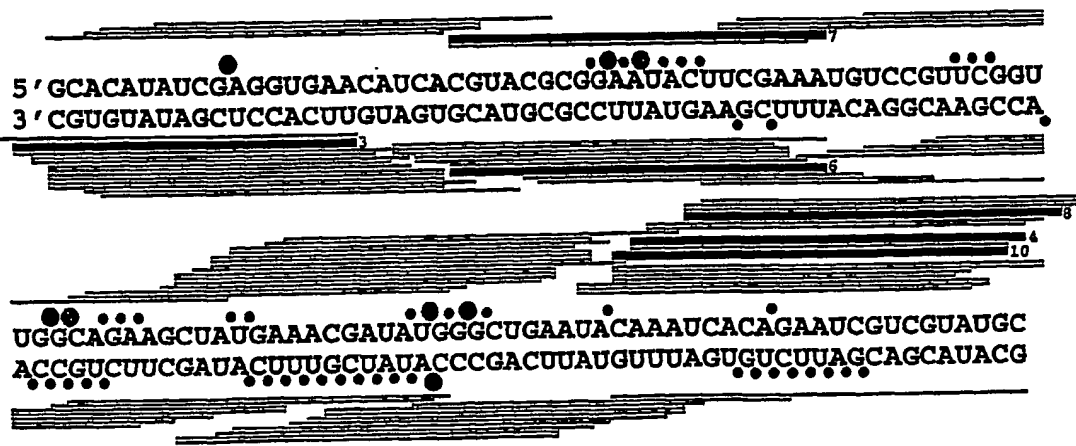


FIGURE 4A

5/23

39 bp ds*Pp*-luc p13352 bp ds*Pp*-luc p133111 bp ds*Pp*-luc p133

6/23

FIGURE 4B

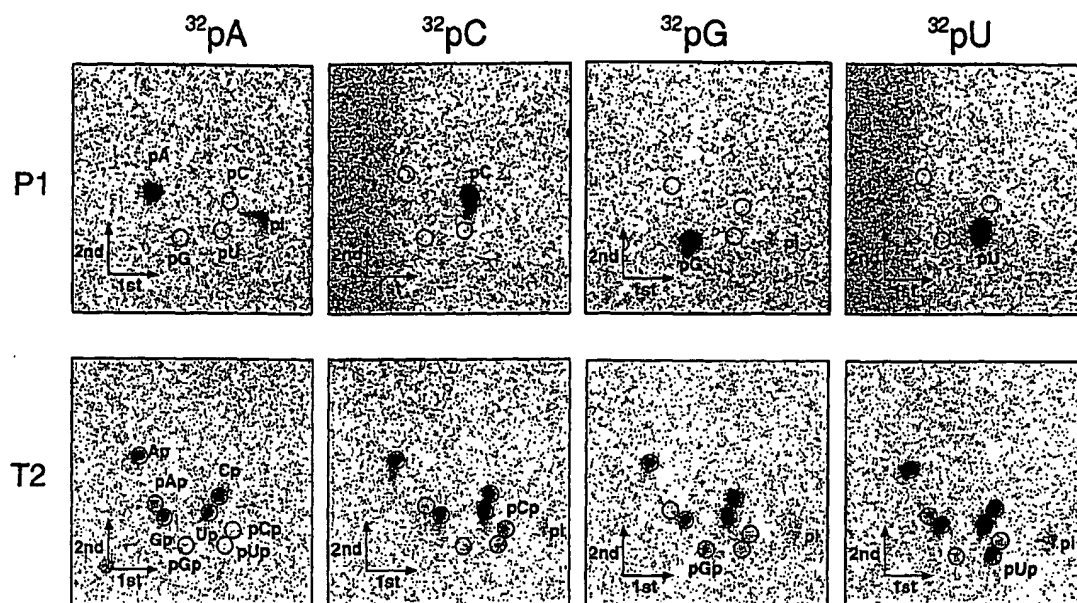
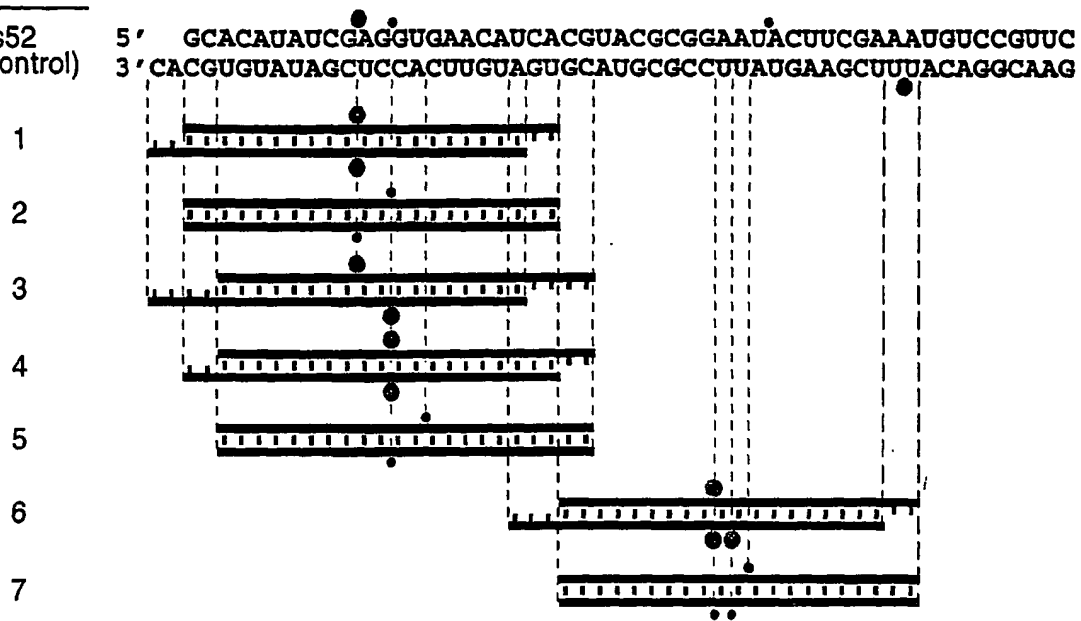


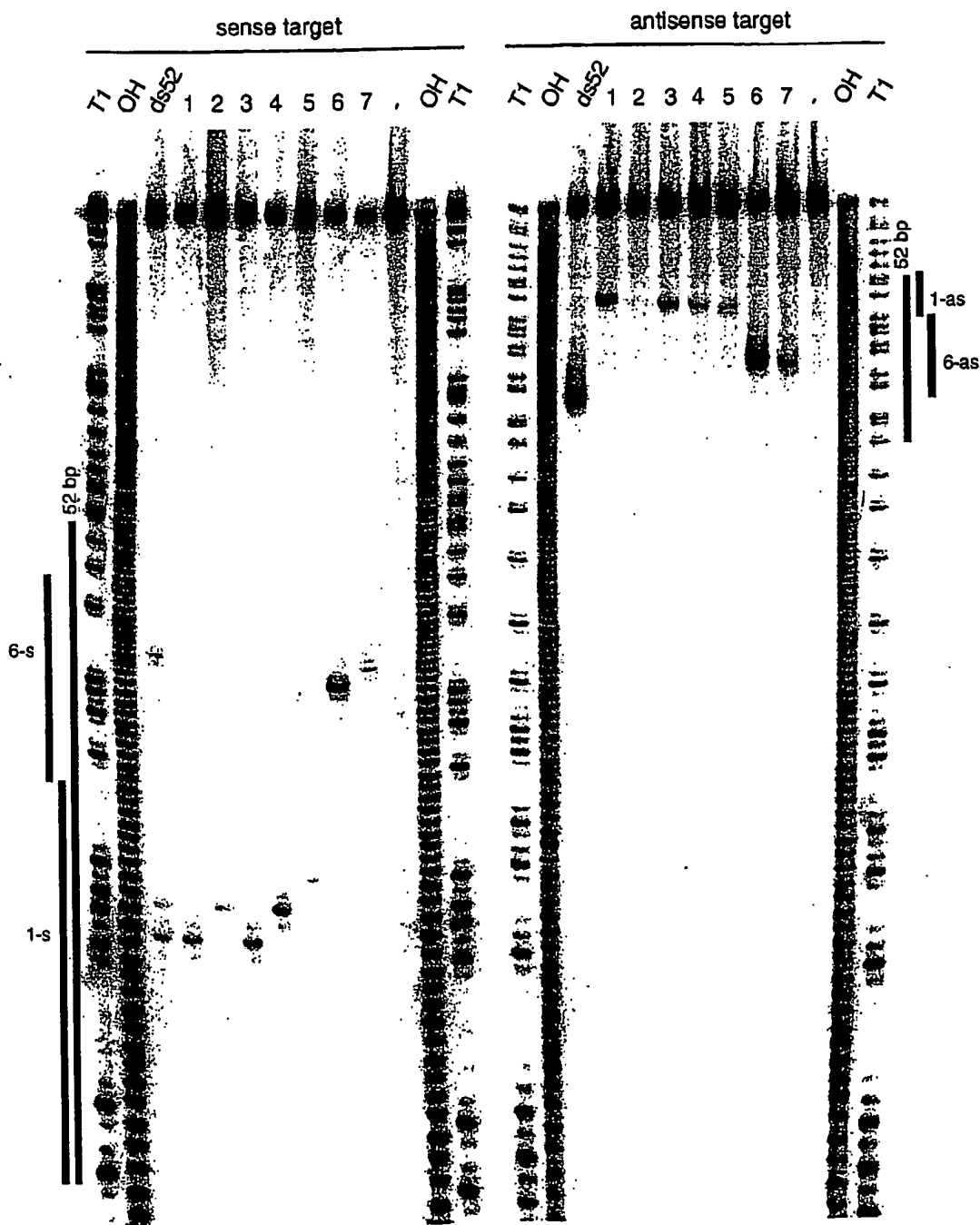
FIGURE 5A

7/23

siRNAs/  
dsRNAds52  
(control)

8/23

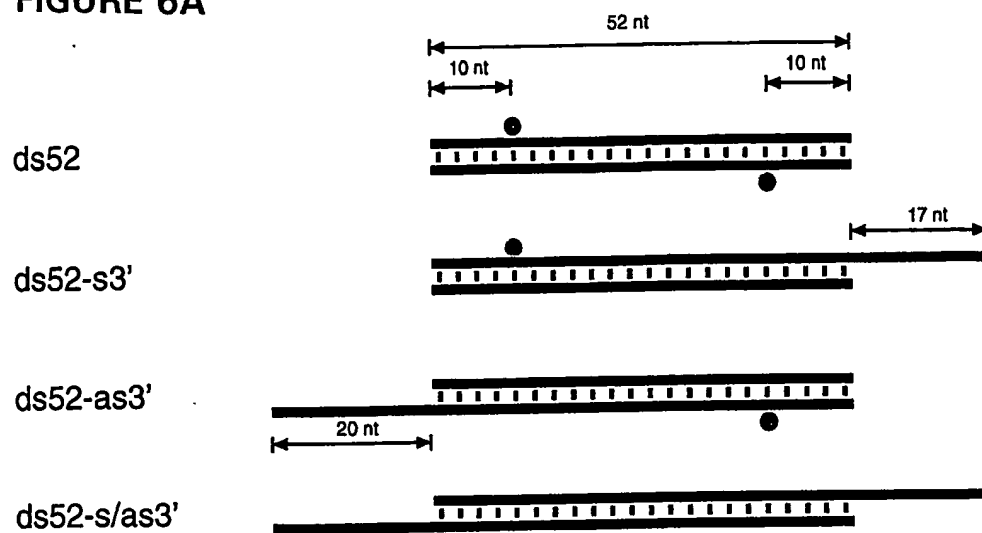
FIGURE 5B





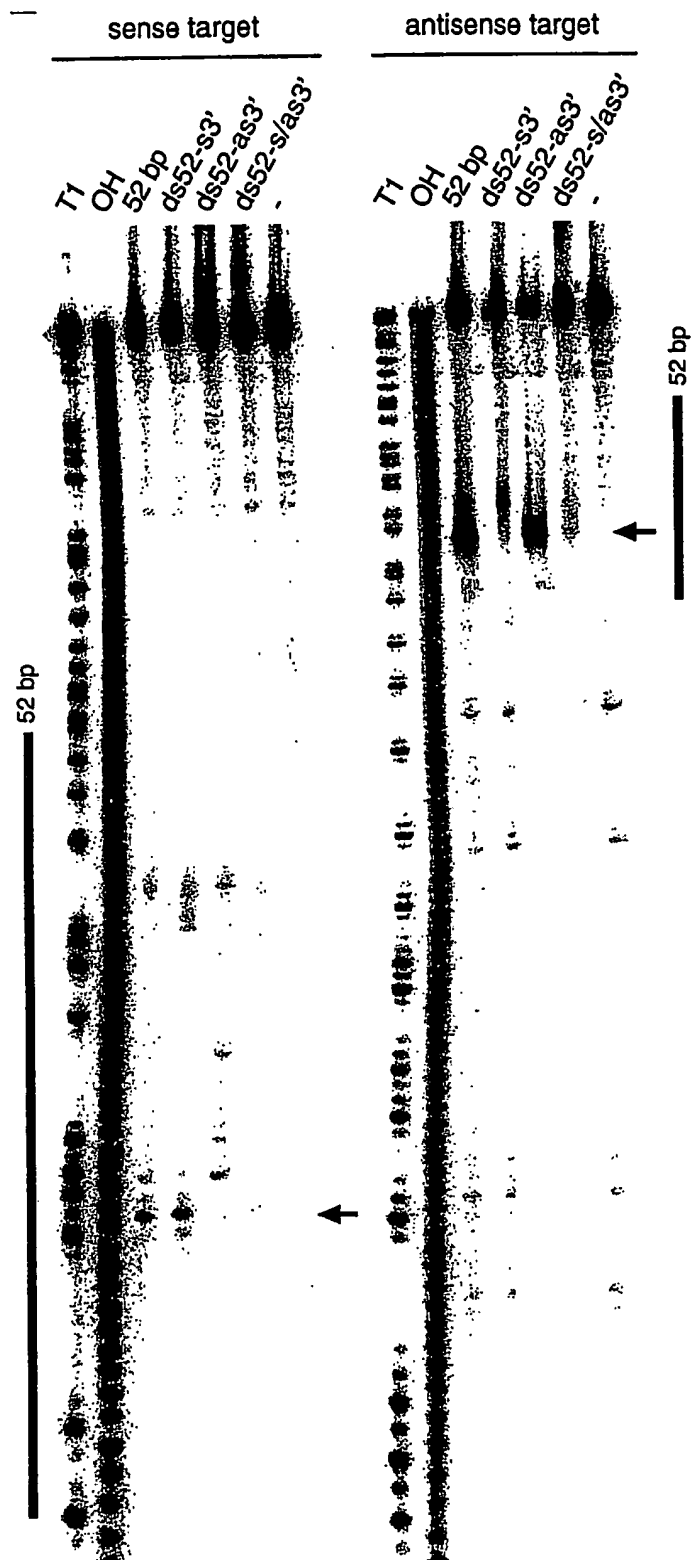
9/23

FIGURE 6A



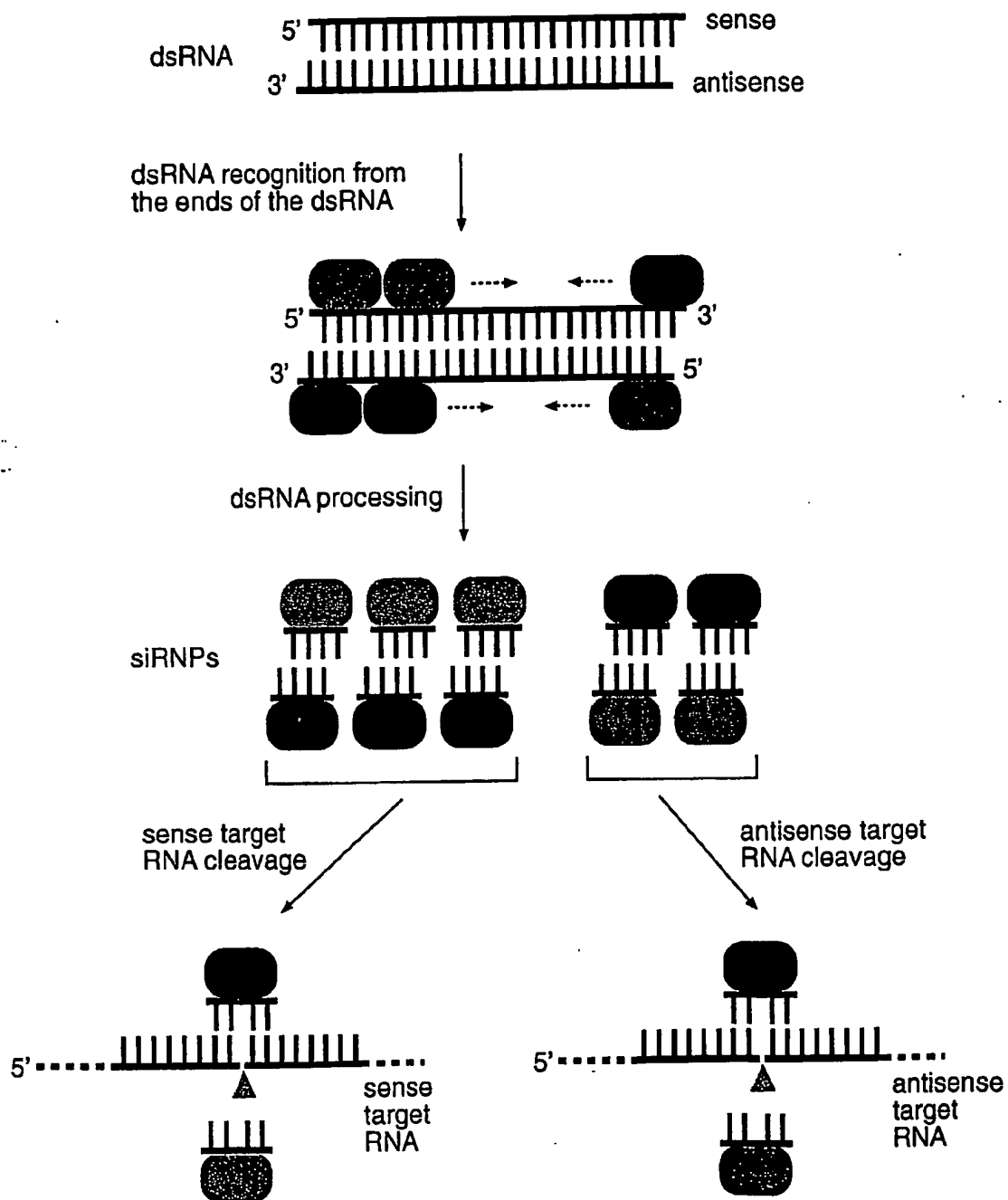
10/23

FIGURE 6B

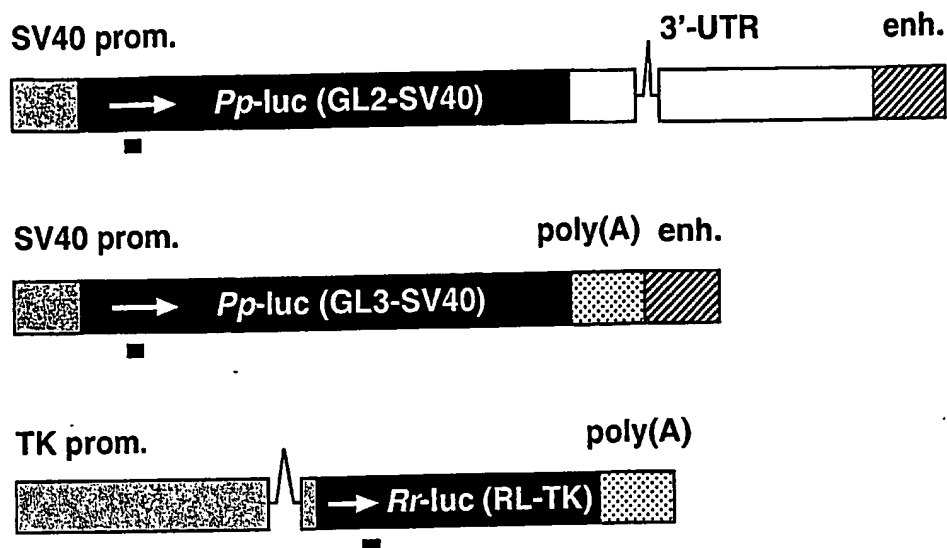


11/23

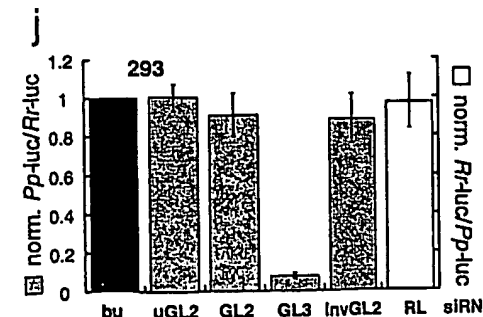
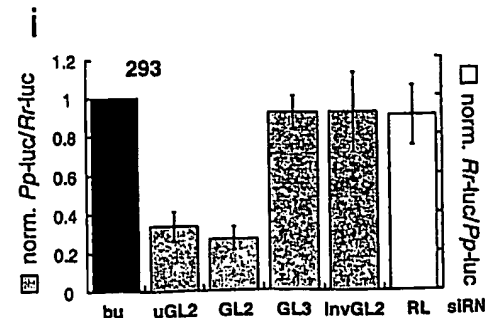
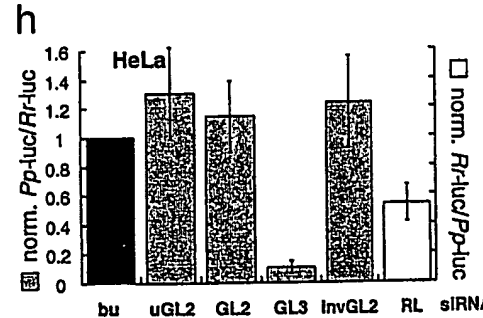
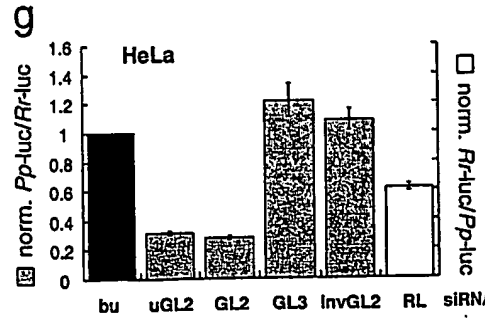
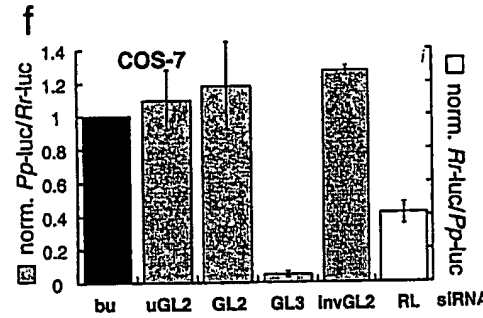
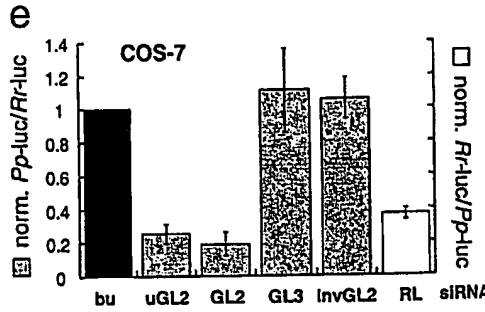
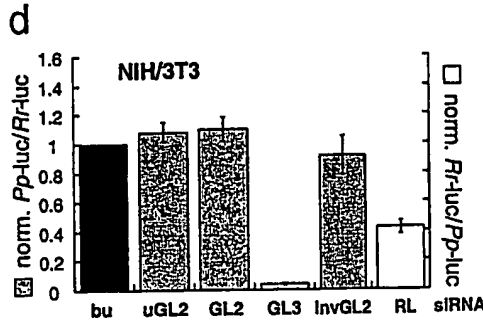
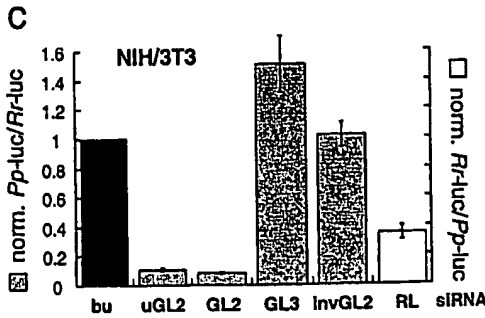
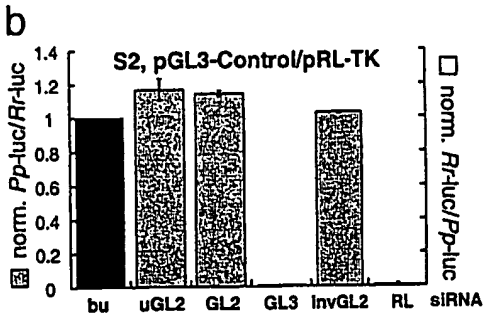
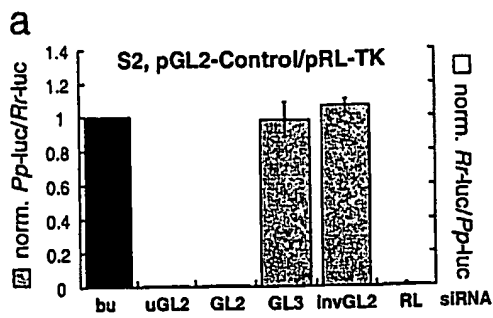
FIGURE 7



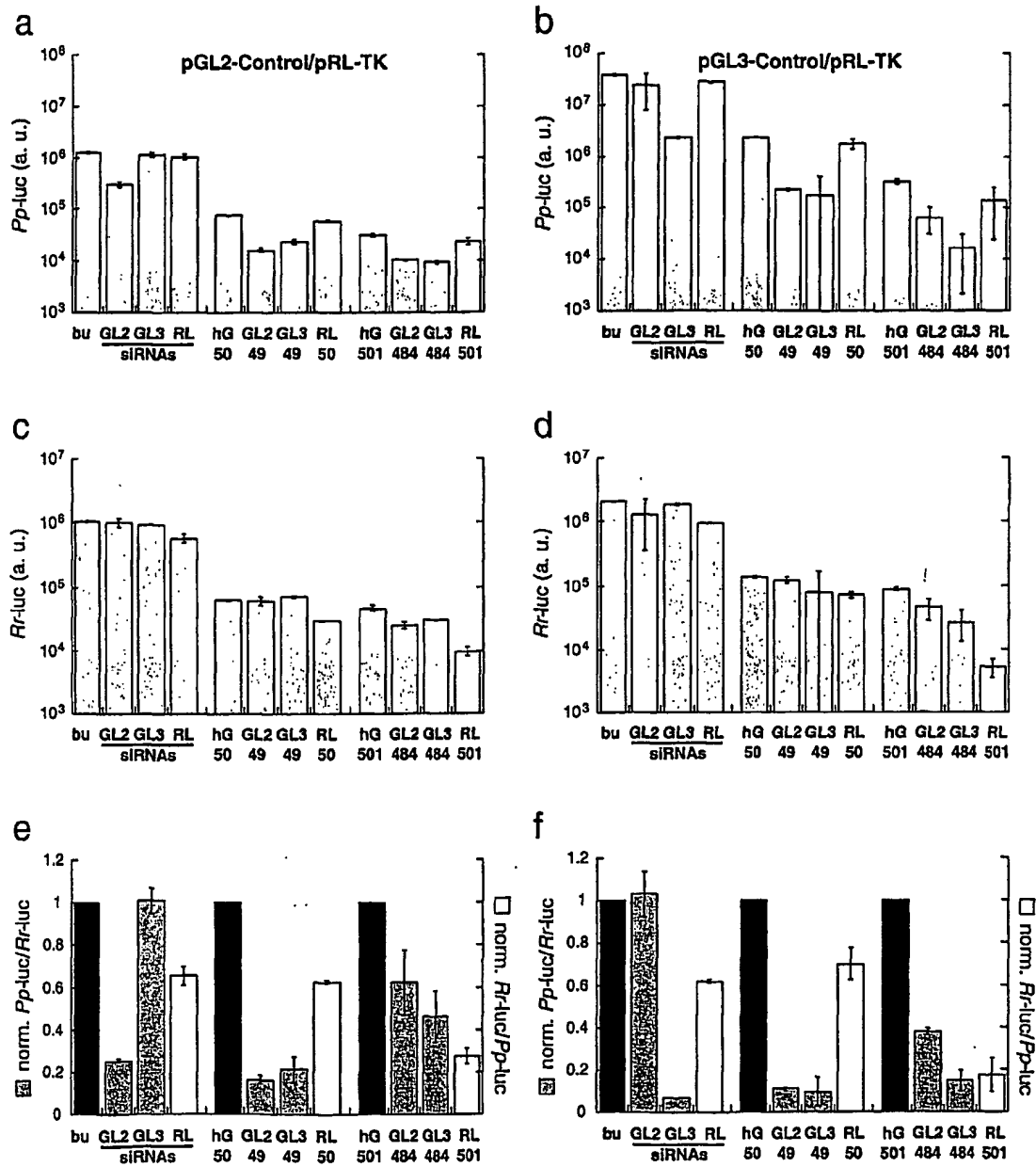
12/23

**a****b**siRNA  
duplex

uGL2	5' CGUACGCGGAUACUUCGAUU UUGCAUGCGCCUUAUGAAGCU 5'
GL2	5' CGUACGCGGAUACUUCGATT TTGCAUGCGCCUUAUGAAGCU 5'
GL3	5' CUUACGCUGAGUACUUCGATT TTGCAUGCGACUCAUGAAGCU 5'
invGL2	5' AGCUUCAUAAGGCGCAUGCTT TTUCGAAGUAUCCGCGUACG 5'
RL	5' AAACAUGCAGAAAUGCUGTT TTUUUGUACGUCUUUUACGAC 5'

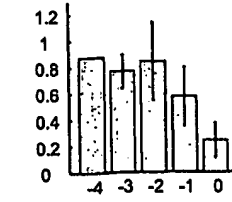
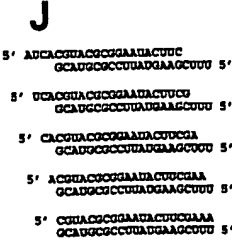
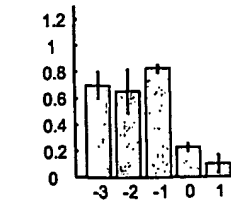
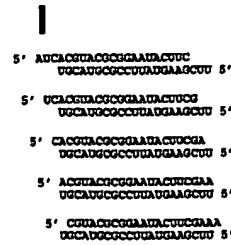
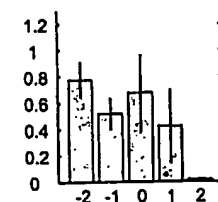
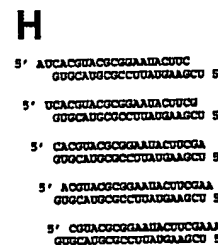
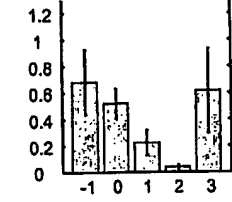
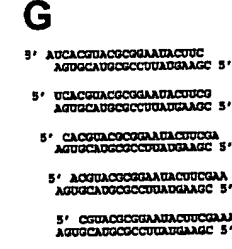
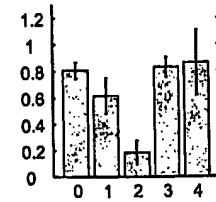
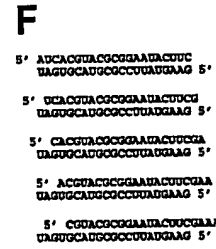
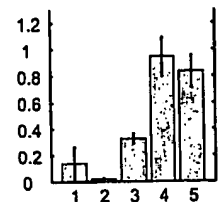
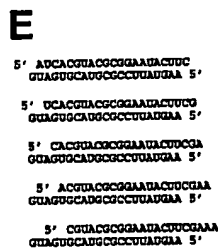
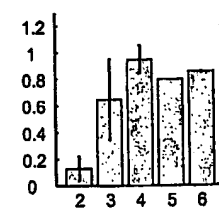
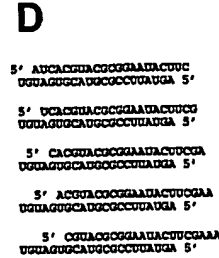
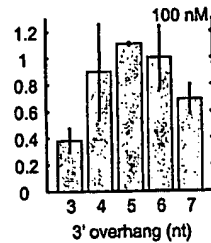
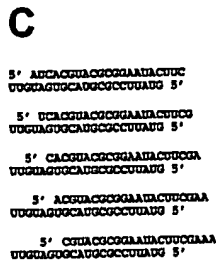
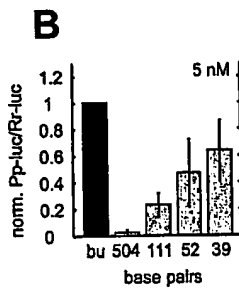
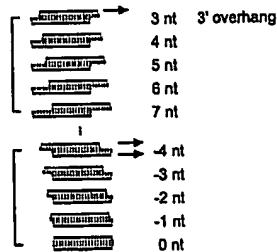
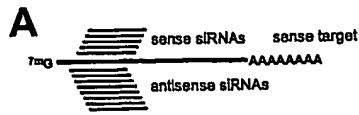


14/23



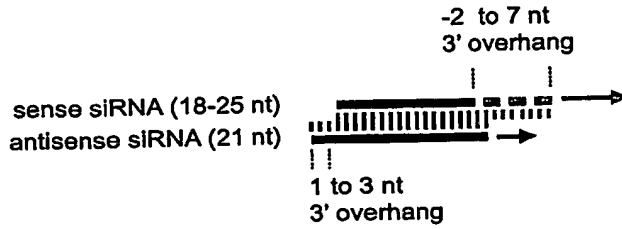
15/23

Fig. 11



16/23

Fig. 12

**A****B**

5' CGUACGCGGAAUACUUCG  
UGCAUGCGCCUUAUGAAGCU 5'

5' CGUACGCGGAAUACUUCGA  
UGCAUGCGCCUUAUGAAGCU 5'

5' CGUACGCGGAAUACUUCGAA  
UGCAUGCGCCUUAUGAAGCU 5'

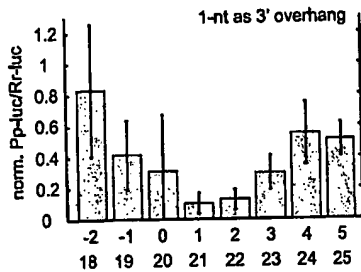
5' CGUACGCGGAAUACUUCGAAA  
UGCAUGCGCCUUAUGAAGCU 5'

5' CGUACGCGGAAUACUUCGAAAU  
UGCAUGCGCCUUAUGAAGCU 5'

5' CGUACGCGGAAUACUUCGAAAU  
UGCAUGCGCCUUAUGAAGCU 5'

5' CGUACGCGGAAUACUUCGAAAU  
UGCAUGCGCCUUAUGAAGCU 5'

5' CGUACGCGGAAUACUUCGAAAU  
UGCAUGCGCCUUAUGAAGCU 5'

**C**

5' CGUACGCGGAAUACUUCG  
GUGCAUGCGCCUUAUGAAGCU 5'

5' CGUACGCGGAAUACUUCGA  
GUGCAUGCGCCUUAUGAAGCU 5'

5' CGUACGCGGAAUACUUCGAA  
GUGCAUGCGCCUUAUGAAGCU 5'

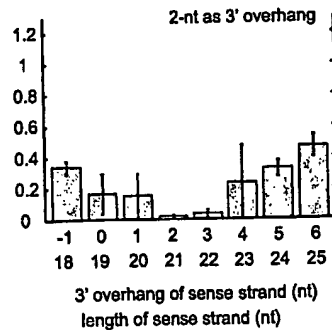
5' CGUACGCGGAAUACUUCGAAA  
GUGCAUGCGCCUUAUGAAGCU 5'

5' CGUACGCGGAAUACUUCGAAAU  
GUGCAUGCGCCUUAUGAAGCU 5'

5' CGUACGCGGAAUACUUCGAAAU  
GUGCAUGCGCCUUAUGAAGCU 5'

5' CGUACGCGGAAUACUUCGAAAU  
GUGCAUGCGCCUUAUGAAGCU 5'

5' CGUACGCGGAAUACUUCGAAAU  
GUGCAUGCGCCUUAUGAAGCU 5'

**D**

5' CGUACGCGGAAUACUUCG  
AGUGCAUGCGCCUUAUGAAGC 5'

5' CGUACGCGGAAUACUUCGA  
AGUGCAUGCGCCUUAUGAAGC 5'

5' CGUACGCGGAAUACUUCGAA  
AGUGCAUGCGCCUUAUGAAGC 5'

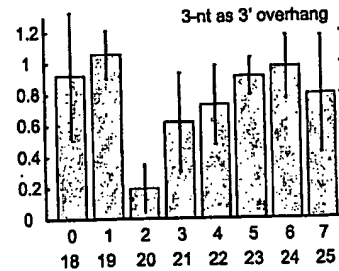
5' CGUACGCGGAAUACUUCGAAA  
AGUGCAUGCGCCUUAUGAAGC 5'

5' CGUACGCGGAAUACUUCGAAAU  
AGUGCAUGCGCCUUAUGAAGC 5'

5' CGUACGCGGAAUACUUCGAAAU  
AGUGCAUGCGCCUUAUGAAGC 5'

5' CGUACGCGGAAUACUUCGAAAU  
AGUGCAUGCGCCUUAUGAAGC 5'

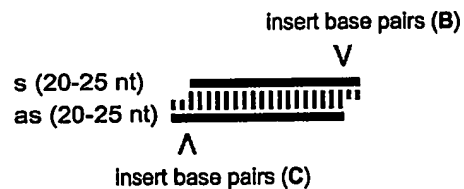
5' CGUACGCGGAAUACUUCGAAAU  
AGUGCAUGCGCCUUAUGAAGC 5'





17/23

Fig. 13

**A****B**

5' CGUACGCGGAUACUUCGAAA  
GUGCAUGCGCCUUAUGAAGC 5'

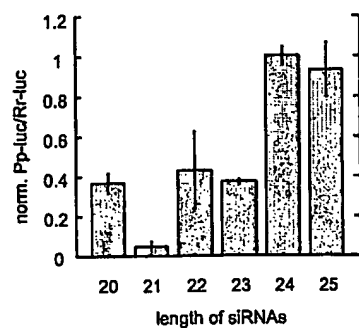
5' CGUACGCGGAUACUUCGAAA  
GUGCAUGCGCCUUAUGAAGC 5'

5' CGUACGCGGAUACUUCGAAA  
GUGCAUGCGCCUUAUGAAGC 5'

5' CGUACGCGGAUACUUCGAAA  
GUGCAUGCGCCUUAUGAAGC 5'

5' CGUACGCGGAUACUUCGAAA  
GUGCAUGCGCCUUAUGAAGC 5'

5' CGUACGCGGAUACUUCGAAA  
GUGCAUGCGCCUUAUGAAGC 5'

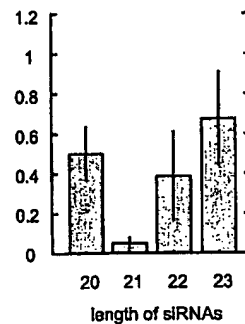
**C**

5' GUACGCGGAUACUUCGAAA  
UGCAUGCGCCUUAUGAAGC 5'

5' CGUACGCGGAUACUUCGAAA  
GUGCAUGCGCCUUAUGAAGC 5'

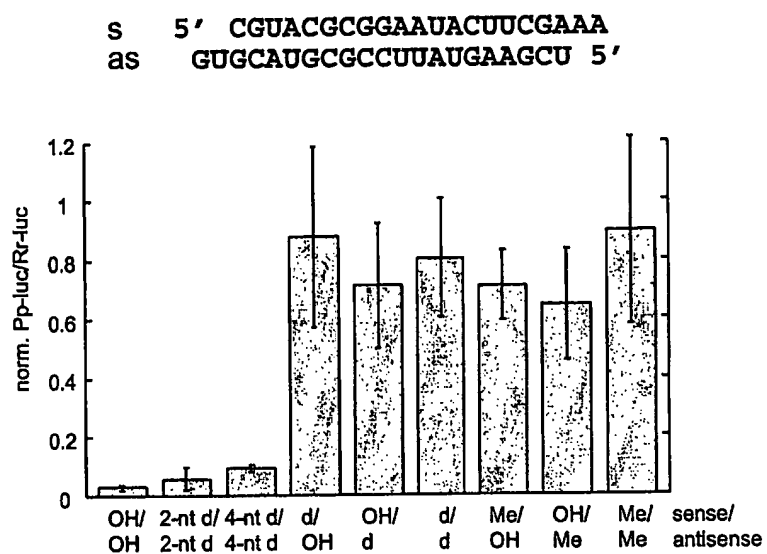
5' ACGUACGCGGAUACUUCGAAA  
AGUGCAUGCGCCUUAUGAAGC 5'

5' CACGUACGCGGAUACUUCGAAA  
UAGUGCAUGCGCCUUAUGAAGC 5'



18/23

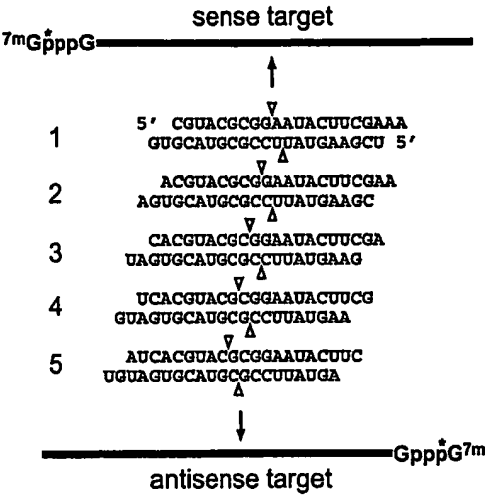
Fig. 14



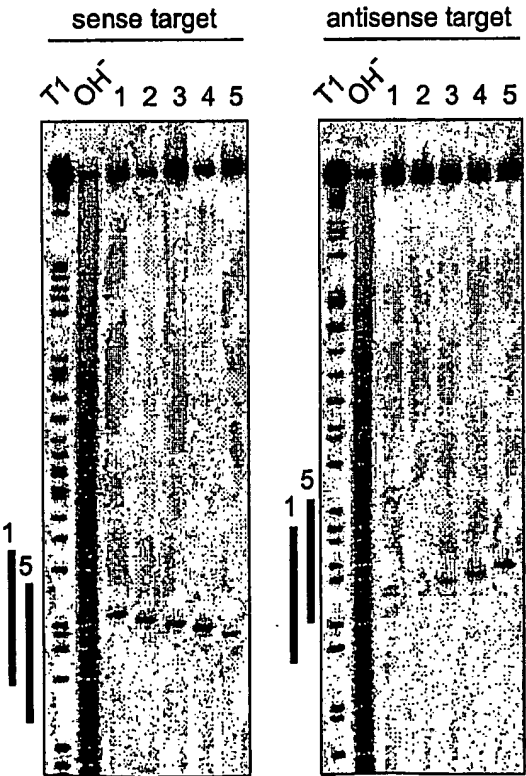
19/23

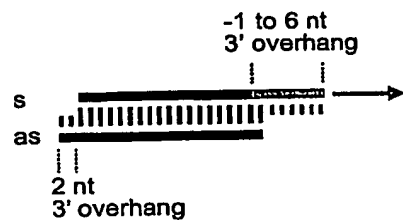
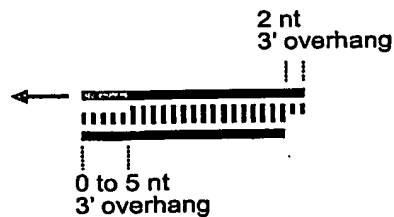
Fig. 15

**A**



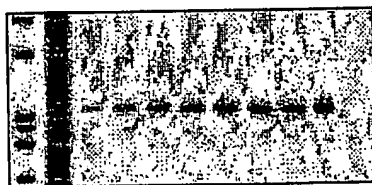
**B**



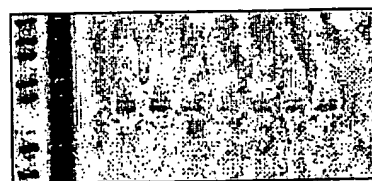
**A****B****C**

5' CGUACGCGGAUACUUCG  
GUGCAUGCGCCUUAUGAAGCU 5'  
5' CGUACGCGGAUACUUCGA  
GUGCAUGCGCCUUAUGAAGCU 5'  
5' CGUACGCGGAUACUUCGAA  
GUGCAUGCGCCUUAUGAAGCU 5'  
5' CGUACGCGGAUACUUCGAAA  
GUGCAUGCGCCUUAUGAAGCU 5'  
5' CGUACGCGGAUACUUCGAAAU  
GUGCAUGCGCCUUAUGAAGCU 5'  
5' CGUACGCGGAUACUUCGAAUUG  
GUGCAUGCGCCUUAUGAAGCU 5'  
5' CGUACGCGGAUACUUCGAAUUGU  
GUGCAUGCGCCUUAUGAAGCU 5'  
5' CGUACGCGGAUACUUCGAAUUGUC  
GUGCAUGCGCCUUAUGAAGCU 5'

18 19 20 21 22 23 24 25 -



sense  
target



antisense  
target

**D**

5' ACGCGGAUACUUCGAAA  
GUGCAUGCGCCUUAUGAAGCU 5'  
5' UACGCGGAUACUUCGAAA  
GUGCAUGCGCCUUAUGAAGCU 5'  
5' GUACGCGGAUACUUCGAAA  
GUGCAUGCGCCUUAUGAAGCU 5'  
5' CGUACGCGGAUACUUCGAAA  
GUGCAUGCGCCUUAUGAAGCU 5'  
5' ACGUACGCGGAUACUUCGAAA  
GUGCAUGCGCCUUAUGAAGCU 5'  
5' CACGUACGCGGAUACUUCGAAA  
GUGCAUGCGCCUUAUGAAGCU 5'

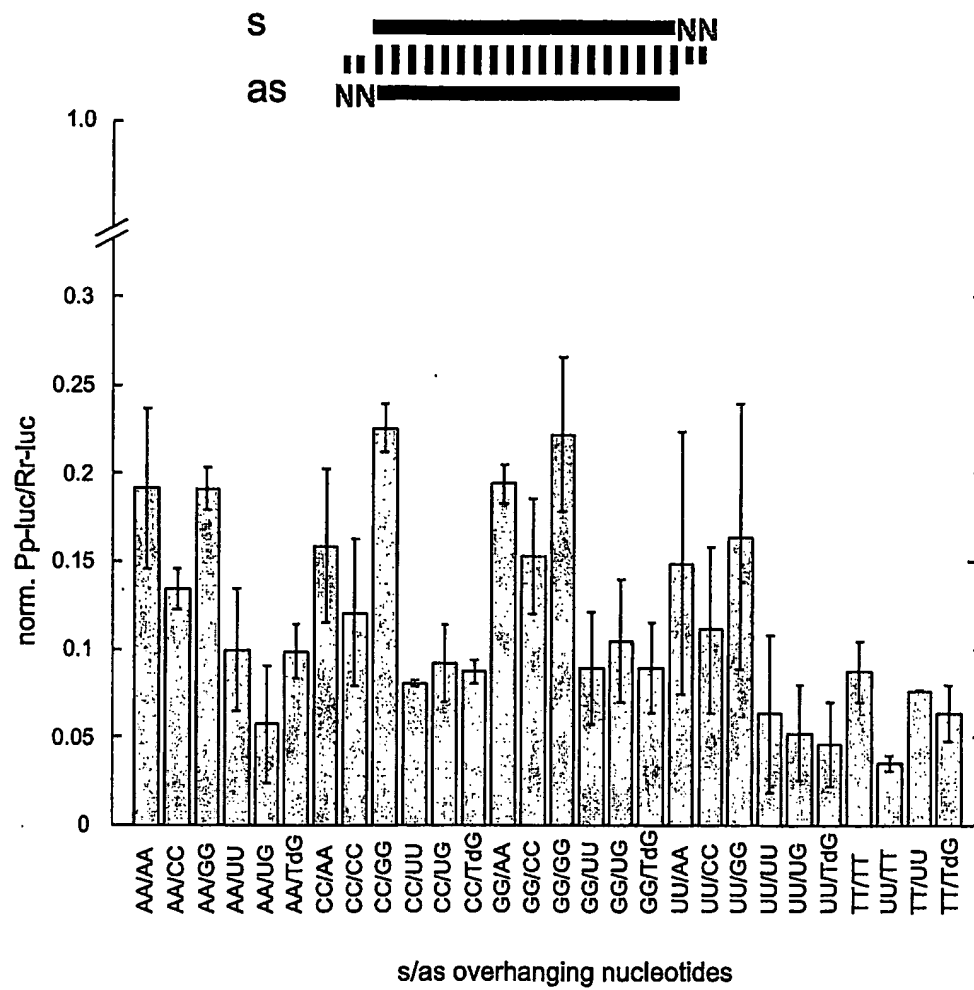
18 19 20 21 22 23



sense  
target



antisense  
target



ref 5' CGUACGCGGAUACUUCGATT  
TTGCAUGCGCCUUAUGAAGCU 5'

1 5' AUGCGCGGAUACUUCGATT  
TTUACGCGCCUUAUGAAGCU 5'

2 5' CGUACGCGGAUACUUCGATT  
TTGCAUGCGCCUUAUGAAGCU 5'

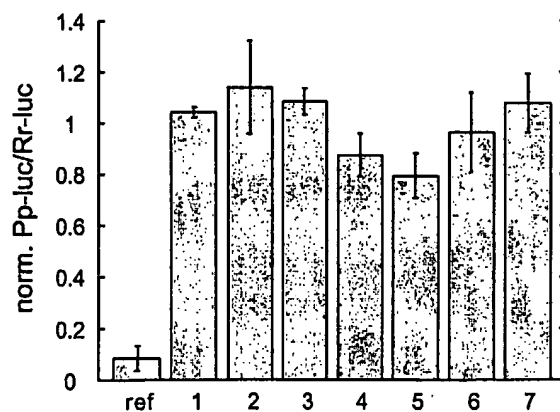
3 5' CGUACGCGGAUACUUCGATT  
TTGCAUGCGCCUUAUGAAGCU 5'

4 5' CGUACGCGGAUACUUCGATT  
TTGCAUGCGCCUUAUGAAGCU 5'

5 5' CGUACGCGGAUACUUCGATT  
TTGCAUGCGCCUUAUGAAGCU 5'

6 5' CGUACGCGGAUACUUCGATT  
TTGCAUGCGCCUUAUGAAGCU 5'

7 5' CGUACGCGGAUACUUCGATT  
TTGCAUGCGCCUUAUGAAGCU 5'



**A**

5' CGUACGCGGAAUACUUCGAA  
GUGCAUGCGCCUUAUGAAGC 5'

5' CGUACGCGGAAUACUUCGAAA  
GUGCAUGCGCCUUAUGAAGCU 5'

5' CGUACGCGGAAUACUUCGATT  
TTGCAUGCGCCUUAUGAAGCU 5'

5' CGUACGCGGAAUACUUCGAAA  
GUGCAUGCGCCUUAUGAAGCUU 5'

5' CGUACGCGGAAUACUUCGAAAUG  
GUGCAUGCGCCUUAUGAAGCUUU 5'

5' CGUACGCGGAAUACUUCGAAAUGU  
GUGCAUGCGCCUUAUGAAGCUUUA 5'

5' CGUACGCGGAAUACUUCGAAAUGUC  
GUGCAUGCGCCUUAUGAAGCUUUAAC 5'

**B**

5' GUACGCGGAAUACUUCGAAA  
UGCAUGCGCCUUAUGAAGCU 5'

5' CGUACGCGGAAUACUUCGAAA  
GUGCAUGCGCCUUAUGAAGCU 5'

5' ACGUACGCGGAAUACUUCGAAA  
AGUGCAUGCGCCUUAUGAAGCU 5'

5' CACGUACGCGGAAUACUUCGAAA  
UAGUGCAUGCGCCUUAUGAAGCU 5'

